CONTAMINACION DE LAS PLAYAS • FAUNA PLEISTOCENICA

INVESTIGACION y CIENCIA



Edición española de SCIENTIFIC AMERICAN

El origen de la vida

- Hacia una anestesia más segura
- De las partículas a la cosmología
- Telescopios de espejo líquido
- Lepidosaurios en movimiento



AGOSTO 2007
6,00 EUROS



Edición española de
SCIENTIFIC AMERICAN

Agosto de 2007

Número 371

3
HACE...
50, 100 y 150 años.

4
APUNTES
Atmósfera...
Matemáticas...
Química...
Biofísica...
Botánica...
Psicología.

34
CIENCIA Y SOCIEDAD
Autoorganización de la vegetación...
Comunicación celular.



38
DE CERCA
Biodiversidad en las salinas,
por Alícia Duró y Francisco Amat





8
Contaminación de las playas
Michel A. Mallin

El crecimiento urbanístico incontrolado a lo largo de las costas de Estados Unidos provoca un aumento de la contaminación por microorganismos patógenos.

26
Hacia una anestesia más segura
Beverley A. Orser

La investigación sobre la potencia y el riesgo de la anestesia general abre nuevas vías para el desarrollo de fármacos selectivos, más seguros y sin efectos secundarios.



40
De las partículas a la cosmología
David Kaiser

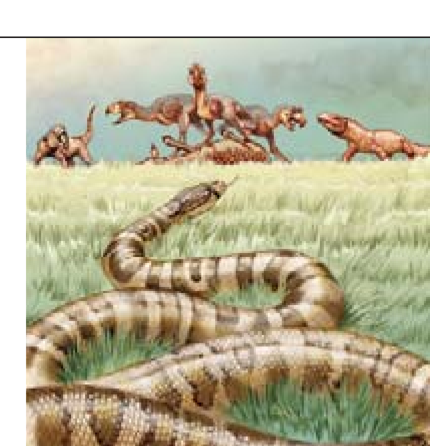
La historia de la cosmología de partículas, nueva rama de la física que intenta dilucidar los orígenes del universo, demuestra que la ciencia puede beneficiarse con los cambios de rumbo.

48
El dilema del viajero
Kaushik Basu

En este sencillo juego, los participantes rechazan sistemáticamente la opción tenida por racional, pero al actuar de forma ilógica acaban cosechando una remuneración mayor. Tal resultado exige un razonamiento formal de nuevo tipo.

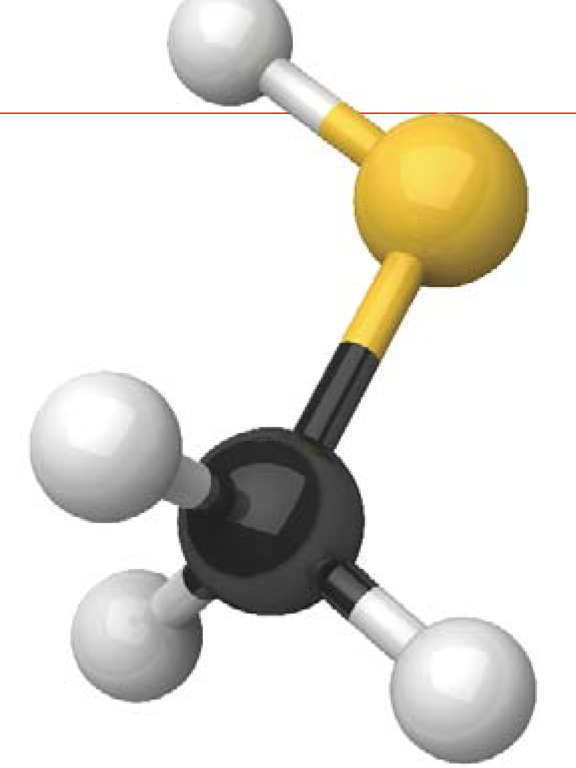
55
Lepidosaurios en movimiento: de los lagartos a las serpientes
Sebastián Apesteguía

"Marinistas" y "terrestristas" sostienen ideas diferentes acerca de la historia de las serpientes: ¿se originaron en el mar o en tierra? ¿En el Cretácico o en el Jurásico?



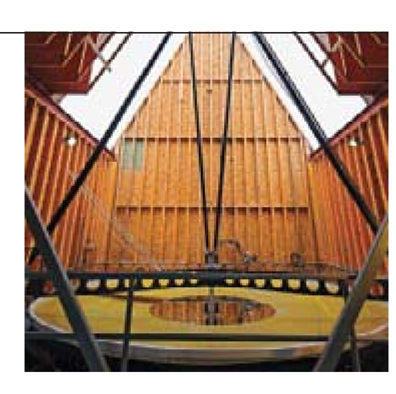
18
El origen de la vida
Robert Shapiro

La aparición repentina de una macromolécula autorreplicante como el ARN era extremadamente improbable. Los iniciadores de la vida habrían sido entramados de reacciones químicas impulsados por una fuente de energía.



62
Telescopios de espejo líquido
Paul Hickson

Una vieja idea para captar imágenes astronómicas está renaciendo gracias a los avances técnicos.

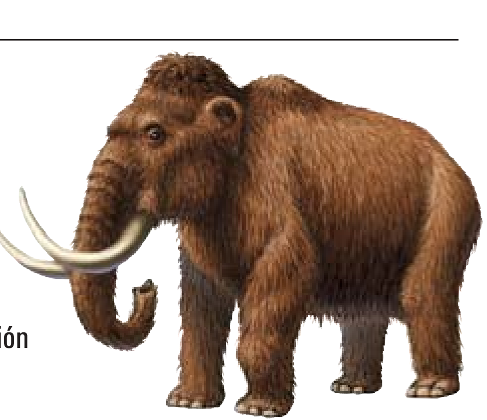


70
Supresión de atascos en las redes de comunicaciones
Michelle Effros, Ralf Koetter y Muriel Médard

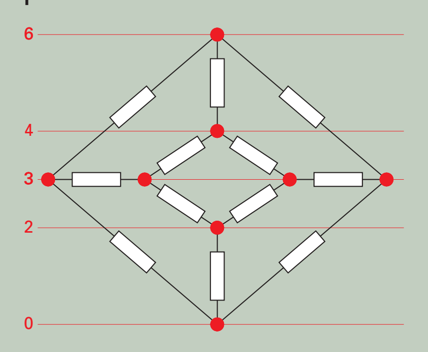
Un nuevo método de transmisión de datos, la “codificación de red”, podría elevar mucho la eficiencia de las redes de comunicaciones. Su idea central es sumamente curiosa: el envío de indicios sobre los mensajes puede resultar más útil que enviar los propios mensajes.

78
Recuperación de la fauna pleistocénica
C. Josh Donlan

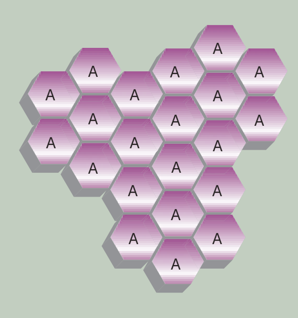
Un proyecto para la recuperación de animales que desaparecieron de Norteamérica hace 13.000 años ofrece un programa optimista para la conservación en el siglo XXI.



87
CURIOSIDADES DE LA FÍSICA
Juegos con resistencias, por Norbert Treitz



90
JUEGOS MATEMÁTICOS
Móviles y vectores, por Juan M.R. Parrondo



92
IDEAS APLICADAS
Reconocimiento de caracteres, por Mark Fischetti



94
LIBROS
Renacimiento.





DIRECTOR GENERAL José M.ª Valderas Gallardo
DIRECTORA FINANCIERA Pilar Bronchal Garfella
EDICIONES Juan Pedro Campos Gómez
Laia Torres Casas
PRODUCCIÓN M.ª Cruz Iglesias Capón
Albert Marín Garau
SECRETARÍA Purificación Mayoral Martínez
ADMINISTRACIÓN Victoria Andrés Laiglesia
SUSCRIPCIONES Concepción Orenes Delgado
Olga Blanco Romero
EDITA Prensa Científica, S.A. Muntaner, 339 pral. 1.ª
08021 Barcelona (España)
Teléfono 934 143 344 Telefax 934 145 413
www.investigacionyciencia.es

SCIENTIFIC AMERICAN

EDITOR IN CHIEF John Rennie
EXECUTIVE EDITOR Mariette DiChristina
MANAGING EDITOR Ricki L. Rusting
CHIEF NEWS EDITOR Philip M. Yam
SENIOR WRITER Gary Stix
SENIOR EDITOR Michelle Press
EDITORS Mark Alpert, Steven Ashley, Graham P. Collins, Mark Fischetti, Steve Mirsky, George Musser y Christine Soares
CONTRIBUTING EDITORS W. Wayt Gibbs, Marguerite Holloway, Michael Shermer, Sarah Simpson
PRODUCTION EDITOR Richard Hunt

CHAIRMAN Brian Napack
VICE PRESIDENT AND MANAGING DIRECTOR, INTERNATIONAL Dean Sanderson
VICE PRESIDENT Frances Newburg
GENERAL MANAGER Michael Florek

DISTRIBUCION

para España:

LOGISTA, S. A.
Pol. Ind. Polvoranca
Trigo, 39, Edif. 2
28914 Leganés (Madrid)
Teléfono 914 819 800

para los restantes países:

Prensa Científica, S. A.
Muntaner, 339 pral. 1.ª
08021 Barcelona

PUBLICIDAD

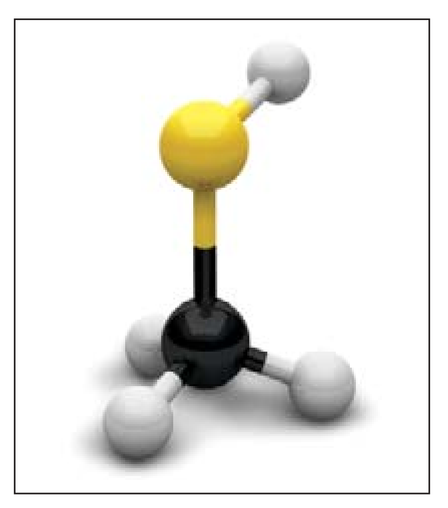
Madrid:
VdS Comunicación
Julia Domínguez
Dr. Fleming, 56 - 3.º dcha.
28036 Madrid
Tel. y fax 913 591 965
Móvil 649 879 433

Cataluña:
QUERALTO COMUNICACION
Julián Queraltó
Sant Antoni M.ª Claret, 281 4.º 3.ª
08041 Barcelona
Tel. y fax 933 524 532
Móvil 629 555 703

COLABORADORES DE ESTE NUMERO

Asesoramiento y traducción:

Bruno Moreno: *Apuntes*; Juan Manuel González Mañas: *El origen de la vida*; J. Vilardell: *Hacia una anestesia más segura*, *Hace...*, *Apuntes* e *Ideas aplicadas*; Ramón Pascual: *De las partículas a la cosmología*; Luis Bou: *El dilema del viajero*, *Supresión de atascos en las redes de comunicaciones*; Ramón García López: *Telescopios de espejo líquido*; Joandomènec Ros: *Recuperación de la fauna pleistocénica*; Jürgen Goicoechea: *Curiosidades de la física*



Portada: Ken Eward, BioGrafx

SUSCRIPCIONES

Prensa Científica S. A.
Muntaner, 339 pral. 1.ª
08021 Barcelona (España)
Teléfono 934 143 344
Fax 934 145 413

Precios de suscripción:

	Un año	Dos años
España	65,00 euro	120,00 euro
Resto del mundo	100,00 euro	190,00 euro

Ejemplares sueltos:

El precio de los ejemplares atrasados es el mismo que el de los actuales.



ISSN 0210136X Dep. legal: B. 38.999 – 76

Imprime Rotocayfo-Quebecor, S.A. Ctra. de Caldes, km 3 - 08130 Santa Perpètua de Mogoda (Barcelona)

Printed in Spain - Impreso en España



...cincuenta años

Datos sobre el tráfico. «Los accidentes de automóvil en EE.UU. son ahora el tema de una investigación a gran escala, cual si fueran una epidemia, y en verdad lo son. El Departamento del Ejército, la Asociación Médica Americana y otras organizaciones de primera importancia están estudiando los diversos aspectos del asunto, desde el diseño de las cabinas de peaje hasta la personalidad de los conductores de camión. Entre otros descubrimientos a destacar recordemos que los sedantes y los tranquilizantes embotan la destreza de los conductores y que los efectos peligrosos del alcohol pueden prolongarse hasta 18 horas tras la ingestión del mismo, por mucho café que se tome.»

Ratones estresados. «No hay duda de que una vida masificada necesita de fármacos tranquilizantes, al menos en el caso de los ratones. Dos colaboradores de la facultad de medicina de la Universidad Johns Hopkins encerraron en botes varios grupos de ratones. A la mitad de los grupos les suministraron Miltown (un ansiolítico primitivo); a la otra mitad, nada. Al cabo de media hora, inyectaron a todos los grupos una dosis letal de anfetaminas. Al final del experimento, los ratones no tranquilizados habían muerto todos; los tranquilizados, en cambio, ni siquiera habían empezado a mostrar dificultades respiratorias.»

...cien años

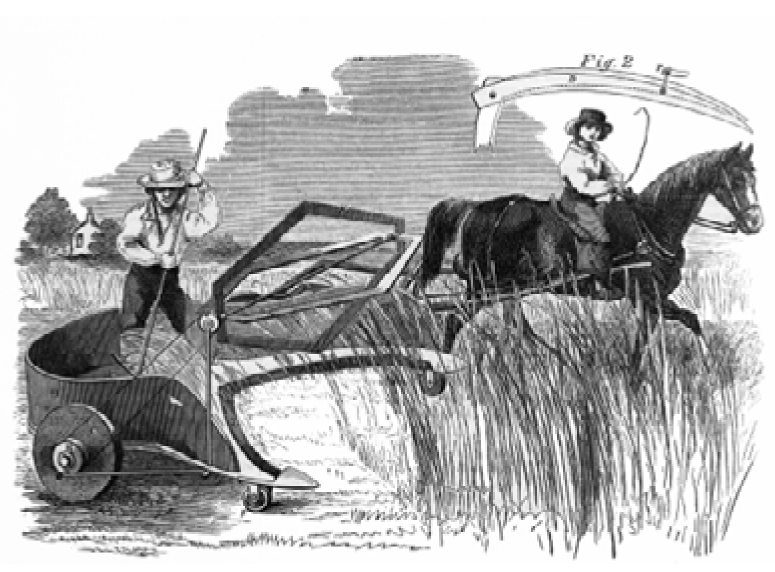
El vapor correo *Lusitania*. «En una prueba preliminar de velocidad, el nuevo trasatlántico de turbinas ‘Lusitania’, de Cunard Line, alcanzó fácilmente los 25 nudos, y ello pese a tener los fondos ‘cubiertos de una gruesa capa del fango, químicamente saturado, del río Clyde’. Las características estructurales del buque son novedosas y de un interés notable, dadas sus colosales proporciones. El doble fondo, que abarca toda la longitud de la nave, tiene una altura de un metro y medio. Posee un total de cinco cubiertas; el casco está dividido en 175 compartimentos estancos independientes, lo cual, desde luego, respalda la afirmación de que ningún desastre natural ordinario sería capaz de hundirlo.”» (*Nota de la redacción*: En 1915 el *Lusitania* fue torpedeado por un submarino alemán y se hundió en 18 minutos.)

Desintegración radiactiva. «Sir William Ramsay ha hecho en fecha reciente un anuncio que, al proceder de autoridad tan reputada, debe acogerse con respeto. Afirma que tras largas experimentaciones con el efecto de varias combinaciones puestas en contacto con la emanación del radio (radón), ha observado que los compuestos de cobre se transmutan, o ‘degradan’, según sus propias palabras, en litio. Después de tratar con la emanación una solución de fosfato de cobre y eliminar luego el cobre, el espectro del residuo muestra la raya roja del litio. De confirmarse ese descubrimiento, debería considerarse la más brillante revelación de esta era de la radiactividad.» (*Nota de la redacción*: El cobre no se desintegra en litio; Marie Curie y Ellen Gleditsch sugirieron en 1908 que el litio procedía del vidrio que confinaba el experimento.)

...ciento cincuenta años

Labranza. «Nuestra ilustración corresponde a una vista en perspectiva de la cosechadora invención de S. Gumaer, de Chicago (Illinois). Cuando la máquina es arrastrada hacia delante por los caballos, una biela genera un peculiar movimiento de vaivén en las cuchillas, cuyos filos siegan las espigas a la manera de las guadañas. Calcula el señor Gumaer que la máquina puede ofrecerse a los agricultores, lista para su uso, al precio de 65 dólares.»

Eter legal. «En Bélgica se ha recurrido a la eterización como medio para conseguir información judicial. Tras un importante robo, dos hombres fueron detenidos y llevados a juicio. El primero fue condenado a trabajos forzados a perpetuidad, pero el juicio del segundo se aplazó en vista de unas pretendidas mudez e idiocia. Se halló imposible extraer de él siquiera la menor señal de inteligencia; pero, sometido a examen médico, bajo los efectos de la eterización, habló perfectamente y en francés. Fue condenado a diez años de trabajos forzados.»



Mecanización de la agricultura: Un intento temprano de 1857.

Scientific American, vol. XII, n.º 51, 29 de agosto, 1857

APUNTES

OZONO

¿Por qué su vacío en la atmósfera y sobreabundancia en la superficie?

La concentración de ozono atmosférico es relativa. Los niveles peligrosamente altos de O_3 en la capa más baja de la atmósfera, la troposfera, resultarían peligrosamente bajos en la estratosfera, la capa siguiente. El ozono que hay a ras del suelo no basta para llenar el agujero de ozono. Además, el ozono está regulado por procesos locales. Una barrera de temperatura en la frontera entre la troposfera y la estratosfera impide la mezcla a gran escala del ozono entre capas atmosféricas.

El ozono estratosférico protege de los rayos ultravioletas de la luz solar. En cambio, unos niveles altos de ozono troposférico pueden acarrear problemas para la salud humana y perjudicar cosechas y bosques.

El ozono atmosférico se genera y elimina mediante procesos naturales. En la estratosfera, los rayos solares ultravioleta descomponen una fracción del oxígeno molecular (O_2). Los átomos de oxígeno separados se recombinan luego con otra molécula de oxígeno para formar ozono.

Algunos contaminantes industriales, es el caso de los halocarburos, llegan a la estratosfera porque son inactivos en la troposfera. Finalmente, se descomponen en moléculas que agotan, como el monóxido de cloro (ClO), el ozono estratosférico mediante su retrotransformacíon en oxígeno molecular.

Los niveles de ozono estratosférico suelen situarse en torno a las 400 unidades Dobson (DU), la unidad de medida estándar de la concentración de ozono. En la primavera antártica, unas condiciones de frío extremo posibilitan reacciones químicas que generan niveles altos del ClO destructor del ozono. Dentro del agujero de ozono de la Antártica, los niveles pueden caer hasta 85 DU.

De ordinario, en la troposfera no hay más de unas 25 DU; el nivel exacto depende mucho de las condiciones locales. A esas altitudes, la producción natural de ozono resulta ineficiente, pues la intensidad de la luz solar ultravioleta es muy baja. La quema de combustibles fósiles y de biomasa genera grandes cantidades de monóxido de carbono, hidrocarburos y óxidos de nitrógeno; estos gases participan en una serie de reacciones químicas que producen ozono troposférico.

El Protocolo de Montreal prohíbe la producción de halocarburos en todo el mundo; se espera que la capa de ozono estratosférico se recupere en los próximos 50 a 100 años. Se quiere implantar controles de las emisiones que limiten el ozono troposférico a valores inferiores a los prescritos. Esas iniciativas tropiezan con la industrialización global y con el efecto que tienen sobre el ozono troposférico los contaminantes emitidos tanto localmente como en fuentes situadas viento abajo, incluso en otros países o continentes.

—Ross J. Salawitch

NASA

MATEMATICAS

La espuma de la cerveza y los números

Hace cincuenta y cinco años, John von Neumann demostró que el área de una región bidimensional sometida a tensión superficial —es el caso de una burbuja— cambiaba proporcionalmente al número de sus lados. (Con cinco o menos, se encoge; con siete o más, crece; con seis, mantiene su área.) Desde entonces, se ha intentado aplicar el resultado de Von Neumann a cierto caso análogo: el crecimiento tridimensional de granos cristalinos microscópicos. David Srolovitz, de la Universidad Yeshiva, y Robert MacPherson, del Instituto de Estudios Avanzados de Princeton, han deducido la ecuación en cuestión para el cambio de volumen en tres (o más) dimensiones inducido por la tensión; con ese fin, han introducido una magnitud abstracta, el "espesor medio". Como el área o el volumen de una figura, el espesor medio es una medida, aunque no intuitiva, del tamaño de una región en unidades de longitud, sea cual sea la forma que tenga. Si se logra aplicar ese resultado a los agregados cambiantes de cristales o de pompas de espuma, podremos proyectar materiales más fuertes o minimizar la espuma en un vaso de cerveza. Según Srolovitz, la fórmula expresa, en esencia, el cambio que experimentará cada burbuja de la espuma de un vaso de cerveza.

—J. R. Minkel



QUIMICA

Rayadores del diamante

En su calidad de mineral dotado de mayor dureza, el diamante es ideal para cortar rocas y otros materiales resistentes. Pero es caro y se degrada en la mecanización del acero y de otros metales ferrosos en virtud de reacciones que crean carburos de hierro, más blandos. Para cortar acero, la opción alternativa es el nitruro de boro cúbico, que tiene casi la misma dureza y resiste de 40 a 50 gigapascal (GPa) de presión; entre 70 y 100 GPa presenta el diamante. Sin embargo, la fabricación de esa sustancia requiere altas temperaturas (1500 ºC) y presiones extremas (5 GPa), que la encarecen mucho. Un método de fabricación a baja presión puede facilitar un camino menos costoso hacia los materiales extraduros. El equipo de Sarah Tolbert, de la Universidad de California en Los Angeles, perseguía diseñar materiales de ese tipo en vez de llegar a ellos por métodos de ensayo y error. La dureza del diamante y del nitruro de boro dimana de los estrechos enlaces covalentes que mantienen a los átomos constituyentes fuertemente unidos entre sí y con rigidez en todas las direcciones. Ciertos metales "ultraincompresibles" son rígidos en dos dimensiones: resisten la compresión, ya que los electrones de su interior se repelen entre sí, pero son blandos porque sus átomos se disponen en capas que pueden deslizarse unas sobre otras. Tolbert y sus colaboradores encontraron un modo de endurecer estos metales: introdujeron otro elemento, que se unía a esas capas con enlaces covalentes e impedía así que deslizasen, pero sin disminuir apenas su incompresibilidad. En el año 2005, los miembros del equipo combinaron el metal osmio con boro a 1000 ºC y a presión ambiente para fabricar un material que es casi tan incompresible y duro como el diamante. Sustituyendo el osmio por renio, más barato, obtuvieron diboruro de renio, duro (unos 48 GPa) como para rayar el diamante y casi tan incompresible como éste. El método podría conducir a compuestos ultraduros aun más baratos.





La fabricación de materiales ultraduros suele requerir presiones extremas. En cambio, el calor basta para convertir polvo de diboruro de renio (*izquierda*) en una pella (*derecha*) tan dura, que raya el diamante.

OLIVIER LABALETTE *TempSport/Corbis (arriba)*; RICHARD B. KANER *U.C.L.A. (abajo derecha)*; JONATHAN B. LEVINE *U.C.L.A. (abajo izquierda)*

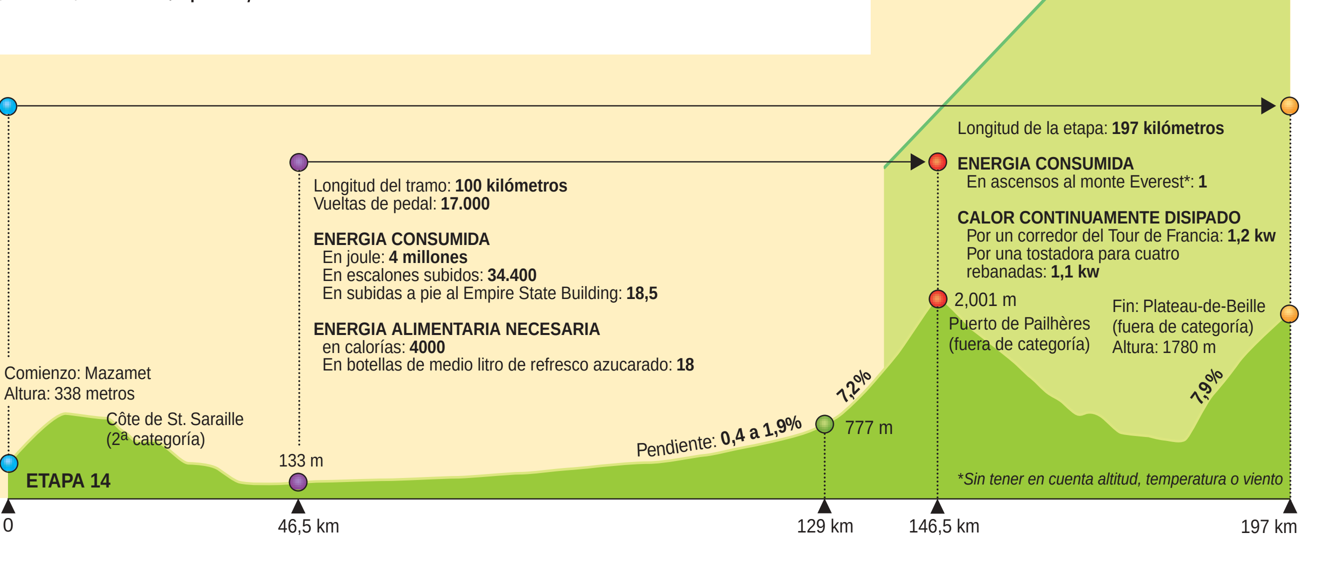


DATOS

El Tour de la energía

Cada día los ciclistas del Tour de Francia consumen cantidades increíbles de energía, especialmente en las etapas de montaña. Hemos pedido al ingeniero David Gordon Wilson, del Instituto de Tecnología de Massachusetts y autor de *La Ciencia del Ciclismo*, que calcule el consumo energético y otros curiosos datos relacionados con la agotadora etapa 14. La dureza de las dos etapas siguientes, también pirenaicas, es similar.

—Mark Fischetti



BIOFISICA
Fotosíntesis cuántica

Un persistente misterio de la fotosíntesis es el de su eficacia al convertir la luz solar en energía química: el rendimiento se aproxima al cien por cien. La clave quizás estribe en la coherencia cuántica, el mismo fenómeno en que se basan los láseres y los semiconductores. En las bacterias púrpura, un conjunto de pigmentos y proteínas absorbe la luz y encauza su energía hacia compuestos químicos. Los elementos que integran este complejo oscilan cuando la luz los excita y esas excitaciones se mantienen sincronizadas mediante vibraciones específicas de las proteínas que conectan los integrantes; recuerda a los impulsos rítmicos con que se mantiene en movimiento un columpio. Tal coherencia hace que el conjunto actúe como una especie de "supermolécula", que se estabiliza de inmediato en la forma que mejor transmite la energía. Esta línea de investigación podría conducir a mejoras de las células solares y otros dispositivos que captan la energía de la luz.

—Charles Q. Choi



KONRAD WOTHE *Minden Pictures*

BOTANICA
La fuerza de las semillas

Una suerte de "músculos" activados por la humedad cambiante ayudan, tal parece, a las semillas del trigo silvestre a alcanzar suelos idóneos donde germinar. La cáscara del grano lleva dos largos apéndices quebradizos, las barbas o aristas, que sostienen el equilibrio de las semillas en su deposición. Un grupo de científicos del Instituto Max Planck de Coloides e Interfases de Golm ha descubierto que las fibrillas de celulosa de las barbas determinan que éstas se expandan con la humedad, con la mutua aproximación consiguiente, y se contraigan y separen con la sequedad. Semejante flexión recuerda el movimiento de las patas de la rana cuando nada, dicen los investigadores. Su objetivo es propulsar las semillas y clavarlas en el suelo. El trigo está optimizado, sostienen, para el ciclo diario de humedad que se presenta en la estación seca (que llega después de la maduración de las semillas) en el Creciente Fértil de Mesopotamia, donde se originó este cultivo.

—Charles Q. Choi



CORTESIA DEL INSTITUTO MAX PLANCK DE COLOIDES E INTERFASES DE GOLM, ALEMANIA

Micrografía electrónica de la sección transversal de una barba, utilizada por las semillas del trigo para enterrarse.

PSICOLOGIA
Las raíces de la animadversión a la ciencia

Conocida es la resistencia en determinados círculos norteamericanos a la teoría de la evolución. En busca de una explicación psicológica del fenómeno, un equipo de investigadores de la Universidad de Yale ha observado que los niños, aun antes de romper a hablar, adquieren ideas de sentido común acerca del mundo físico que pueden persistir hasta la edad adulta y chocar con los descubrimientos de la ciencia. Por ejemplo: como los objetos caen si no los sujetamos, a los niños les cuesta aceptar que el mundo sea redondo; les parece que en el otro lado del mundo los objetos caen abismo abajo. Por otra parte, nociones psicológicas intuitivas llevan a los niños a pensar que todas las cosas están proyectadas en razón de algo: las nubes, para llover. Un supuesto que puede desembocar en un rechazo de la teoría de la evolución. Los psicólogos en cuestión subrayan que, cuando adultos y niños alcanzan de otros el conocimiento, juzgan los temas en función de la autoridad que tales personas les merecen. Ven ahí los autores una posible explicación de la resistencia desmesurada a la ciencia.

—Charles Q. Choi



© *iStockphoto*/OLGA SOLOVEI

FISICA

Atajos dimensionales

¿Aporta pruebas de la teoría un experimento con neutrinos?

El neutrino es un bicho raro en la física de partículas. Sin carga, raramente interacciona con otras partículas y se presenta en tres sabores: electrónico, muónico y táunico; oscila de un sabor a otro mientras se mueve. Durante los últimos cinco años se han estado disparando haces de neutrinos muónicos contra el detector MiniBooNE, un depósito esférico que contiene 800 toneladas de aceite mineral, sito en el Laboratorio Nacional del Acelerador Fermi, en Batavia. Se quería ver cuántas de esas partículas se convertían en neutrinos electrónicos durante el vuelo hacia el detector. Los primeros resultados, anunciados en abril, respaldan el modelo estándar, la teoría vigente en física de partículas, salvo por una anomalía en los datos que no se ha conseguido explicar y que abre la puerta a posibilidades más peculiares. Algunos conjeturan que se debe a un nuevo tipo de neutrino que toma el atajo a través de las dimensiones extra que predice la teoría de cuerdas.

MiniBooNE se creó para contrastar los resultados de un experimento del Laboratorio Nacional de Los Alamos, desarrollado diez años atrás, que apuntaban a la existencia de un cuarto tipo de neutrino. Ese neutrino "estéril", más esquivo que los tres sabores habituales, no estaría sujeto a la fuerza nuclear débil; sólo interaccionaría con otras partículas mediante la gravedad. Puesto que la existencia de neutrinos estériles desafiaría al modelo estándar, se tenía mucho interés en llevar a cabo un experimento similar para confirmar o desechar el hallazgo. Sin embargo, los resultados de MiniBooNE fueron variopintos. En el caso de los neutrinos con energías desde 475 millones de electronvolt hasta tres mil millones de electronvolt, el número de oscilaciones entre sabores se ajustaba bien a las predicciones del modelo estándar; con energías más bajas, se encontró un exceso significativo de neutrinos electrónicos.

Lo que resulta aún más extraño es que tres físicos habían previsto esos resultados. Su trabajo deriva de la teoría de cuerdas, que estipula la existencia de al menos 10 dimensiones para crear una estructura que incorpora tanto la gravedad como la mecánica cuántica. Para explicar por qué no percibimos las otras dimensiones, se supone que las partículas ordinarias de nuestro universo podrían estar confinadas en una "brana" cuatridimensional que flota en un espacio de más dimensiones, como una tira de matamoscas suspendida en el aire. Sólo algunas partículas especiales pueden entrar y salir de la brana, en particular el gravitón (que transmite la fuerza gravitatoria) y el neutrino estéril. En el año 2005, Heinrich Päs, actualmente de la Universidad de Alabama, Sandip Pakvasa, de la Universidad de Hawai, y Thomas J. Weiler, de la Universidad Vanderbilt, propusieron que, si la brana estuviera curvada o microscópicamente deformada, los neutrinos estériles podrían tomar atajos a través del espacio de más dimensiones. Estos atajos influirían en las oscilaciones de sabor, aumentando la probabilidad de una transición para los neutrinos con ciertas energías.



PETER GINTER

Los tubos fotomultiplicadores captan los destellos de luz procedentes de las interacciones de los neutrinos.

Pues bien, los resultados de MiniBooNE concuerdan con las predicciones de Päs, Pakvasa y Weiler. Esta similitud impresionó tanto a algunos investigadores que participaban en el experimento, que enviaron correos electrónicos de felicitación a los tres físicos teóricos. Puesto que no se ha conseguido hasta ahora encontrar pruebas experimentales que demuestren la teoría de cuerdas, la confirmación de la existencia de otras dimensiones constituiría un hallazgo excepcional.

Pero la similitud podría no ser más que una coincidencia. Los investigadores del MiniBooNE están volviendo a comprobar sus datos; intentan determinar si algún efecto imprevisto del entorno o un análisis defectuoso podrían haber alterado el recuento de neutrinos electrones. *—Mark Alpert*

FISICA

Arco iris de láser

Unas partículas semiconductoras, los nanocristales, podrían estar señalando el camino hacia los láseres de colores distintos de los tradicionales rojos y azules que vemos en las lectoras de códigos de barras y reproductores de DVD. Los semiconductores crean la luz láser ofreciendo a los electrones la elección entre dos estados energéticos: alto o bajo. La longitud de onda de la luz emitida depende de la banda prohibida, o diferencia entre las energías de ambos estados, que en los nanocristales disminuye con el tamaño. Víctor Klimov, del Laboratorio Nacional de Los Alamos, y sus colaboradores han rebajado mucho la energía necesaria para hacer láseres de nanocristal; para ello enrollaron seleniuro de cinc en torno a núcleos de sulfuro de cadmio, produciendo cristales de entre cuatro y 20 nanómetros de ancho. Los láseres construidos con estos nanocristales de doble capa podrían ser más baratos y eficientes que los modelos actuales de semiconductores, que constan de capas apiladas hasta un espesor de varias micras.

—J. R. Minkel



SERGEI IVANOV, *Laboratorio Nacional Los Alamos (nanocristales)*

Los nanocristales semiconductores emiten colores que varían según el tamaño de los cristales.



LAS BACTERIAS procedentes de heces animales y humanas amenazan la salud de bañistas y mariscadores. *Escherichia coli* y otros microorganismos patógenos son transportados hasta la costa por las escorrentías pluviales procedentes de zonas residenciales y comerciales, así como por las filtraciones de fosas sépticas mal emplazadas.

CONTAMINACION DE LAS PLAYAS

El crecimiento urbanístico incontrolado a lo largo de las costas de Estados Unidos provoca un aumento de la contaminación por microorganismos patógenos

Michel A. Mallin

El sinuoso e impresionante perfil del litoral estadounidense ejerce un atractivo casi religioso sobre la imaginación de los habitantes del país. Quizá sea Herman Melville quien mejor describe esa atracción en las primeras páginas de *Moby Dick*: "No pararán hasta llegar al límite más extremo de la tierra... Tienen que acercarse al agua todo lo que puedan sin caer dentro de ella".

En los últimos años, millones de estadounidenses han cambiado su residencia para trasladarse a zonas costeras, sobre todo al sudeste, para disfrutar de la naturaleza, un clima suave y una amplia oferta de actividades de ocio. Para nuestro infortunio, el desarrollo rápido y mal planificado está echando a perder, a un ritmo trepidante, tales encantos: crece el número de playas y fondos de asentamiento de mariscos contaminados por microorganismos patógenos procedentes de desechos animales y humanos.

Según un informe reciente del Consejo estadounidense de Defensa de los Recursos Naturales, durante 2004 los estados costeros ordenaron 19.950 días de cierre y avisos de contaminación que afectaban a 1234 zonas de baño, costeras o del interior; es decir, casi una tercera parte de todas las que vigilan regularmente las autoridades sanitarias. El número total de días de playa sujetos a esas medidas fue superior en un 9 por ciento al de 2003 (y a su vez un 50 por ciento superior al total de 2002, aunque este salto tan brusco obedecía también a cambios en las normas de vigilancia federales). El 85 por ciento de los cierres y avisos se debieron a la detección de una concentración excesiva de bacterias fecales en las aguas de la playa.

Las heces animales arrastradas por el agua de escorrentía que generan las tormentas y las de origen humano procedentes del rebosamiento de las alcantarillas o de las infiltraciones de fosas sépticas, incorporan microorganismos que atacan al hígado, infectan el aparato respiratorio y causan trastornos gastrointestinales de consecuencias fatales. Esas enfermedades, que son comunes en países en vías de desarrollo, con sistemas de saneamiento deficientes, aparecen en EE.UU. no como consecuencia de la pobreza sino de un desarrollo insostenible. La construcción desmesurada de viviendas, carreteras, centros comerciales y zonas de estacionamiento ha alterado los sistemas de drenaje naturales de las zonas costeras; las aguas residuales que antaño se infiltraban en los campos o los humedales ensucian hoy las playas y demás zonas litorales.

¿Cómo pueden reducir la contaminación microbiana los estados y comunidades de la costa? Ese tema ha conducido al enfrentamiento entre promotores inmobiliarios y políticos partidarios del crecimiento urbano y las autoridades reguladoras, mariscadores aficionados y profesionales, surfistas, buceadores y ecologistas. Afortunadamente, existen soluciones. Mediante estrategias innovadoras de "crecimiento sostenible" pueden restaurarse las costas contaminadas al propio tiempo que se obtienen beneficios económicos. Dado que el frecuente cierre de las playas amenaza al turismo y causa una devaluación inmobiliaria, la imposición de limitaciones razonables al crecimiento urbanístico de esas áreas aseguraría la economía de la franja litoral, a la vez que protegería la salud de sus habitantes.

Sobreexplotación del litoral

La Administración Nacional Oceánica y Atmosférica (NOAA) señaló en su Informe de Tendencias de la Costa en 2004 que 153 millones de habitantes —el

EMILY HARRISON; FUENTES: EMILY HARRISON, STEF OKASAKI, BEN HERSTON, EYE OF SCIENCE/PHOTO RESEARCHERS, INC., KEVIN FLEMING *Corbis* Y MARY GODLESKI *AP Photo*

53 por ciento de la población estadounidense— residían en condados situados en zonas costeras y en los Grandes Lagos, en una superficie que corresponde sólo al 17 por ciento del área continental del país. Por si fuera poco, se espera que para 2008 esa población haya crecido en siete millones, cifra que se engrosa de forma notable por la afluencia de veraneantes. Extensas zonas boscosas y terrenos agrícolas se han convertido en complejos turísticos, urbanizaciones, centros comerciales, restaurantes, centros de oficinas y polígonos industriales. Para ello, se han desecado humedales y recubierto suelos anteriormente verdes con asfalto, hormigón y otros materiales de construcción.

En el paisaje resultante dominan las superficies impermeables: aparcamientos, carreteras, aceras para peatones, techumbres y terrenos de construcción compactados por maquinaria pesada. Cuando llueve, el agua que fluye sobre esas superficies arrastra heces de animales y otros contaminantes hacia canales de drenaje o colectores de aguas pluviales; la mayoría desembocan en lagos urbanos, arroyos costeros o playas. Si bien las plantas depuradoras eliminan bacterias patógenas y otros contaminantes de las aguas residuales de las alcantarillas, el agua de escorrentías pluviales no suele tratarse. Ya que esos residuos proceden de una zona extensa y no de una fuente concreta, se consideran contaminación no puntual. La Agencia de Protección Ambiental (EPA) declara que ese tipo de contaminación constituye la causa principal de los problemas de calidad del agua en los EE.UU.

Las escorrentías pluviales arrastran fertilizantes, plaguicidas, metales pesados y productos petroquímicos, pero son los microorganismos patógenos (bacterias, virus y protozoos fecales) los que constituyen una seria amenaza para la salud humana. Un solo gramo de heces caninas, por ejemplo, contiene alrededor de 23 millones de bacterias (y bastan sólo 10 bacterias, en el caso de especies o cepas virulentas, para causar una infección). Las cuencas urbanas y suburbanas reciben constantemente un aporte de heces de animales domésticos (perros, gatos) y de la fauna silvestre (ardillas, mapaches). Allí donde existe vegetación, el agua pluvial y los vertidos torrenciales se filtran a través del suelo, que depura el agua de virus y bacterias fecales, además de filtrar otros muchos contaminantes. Las superficies impermeables, en cambio, acumulan la contaminación durante los períodos de sequía; cuando llueve, la dejan correr.

El problema reviste especial gravedad en las costas, ya que los microorganismos contaminan las zonas de baño y los fondos marinos donde medran moluscos. Esos animales se alimentan mediante la filtración de grandes cantidades de agua de mar a través de su cuerpo. Obtienen así algas microscópicas y otros nutrientes; pero también retienen organismos potencialmente dañinos. Si una persona consume moluscos crudos o poco cocidos provenientes de aguas contaminadas por microorganismos fecales, corre el riesgo de contraer gastroenteritis (afección que cursa con vómitos, diarrea y dolor de estómago), amén de otras enfermedades de mayor gravedad.

Resumen/*Contaminación microbiana*

- Debido al espectacular crecimiento de las regiones costeras, gran parte del terreno próximo al mar está impermeabilizado por zonas de estacionamiento, carreteras y aceras. Las escorrentías pluviales arrastran heces animales, plagadas de microorganismos, hacia zanjas de drenaje que vierten las aguas directamente a lagos, arroyos y playas.
- La contaminación por bacterias fecales es la causa principal de las señales de aviso y cierres de playas, que hoy afectan a una tercera parte de las playas estadounidenses vigiladas. Microorganismos patógenos invaden también puertos deportivos, arroyos formados por la pleamar y viveros de mariscos.
- Para combatir la contaminación microbiana, las comunidades costeras deben promover la conservación de las zonas verdes, la instalación de filtros en los sumideros pluviales e impedir la construcción de fosas sépticas en terrenos porosos.

Para proteger a la población, las autoridades deben colocar anuncios, en las zonas donde los fondos están contaminados, que adviertan de la prohibición de recoger almejas, mejillones u ostras. El Servicio de Sanidad Pública estadounidense ha promulgado a escala nacional una norma de seguridad para los viveros de mariscos, que se basa en la medición de la concentración de coliformes fecales, un grupo amplio de microorganismos que medran en los intestinos de personas y animales.

Se prohíbe la recolección de moluscos allí donde la media geométrica del conteo de bacterias en el conjunto de 30 muestras sea mayor que 14 UFC/100 ml (unidades que forman colonias en 100 mililitros de agua). (La media geométrica minimiza el efecto de los valores fuera de margen.) En 1995, la recolección estaba restringida o prohibida en el 31 por ciento de las zonas de cría de moluscos del país. Según la NOAA, la contaminación que con más frecuencia afectaba a los viveros eran los vertidos urbanos.

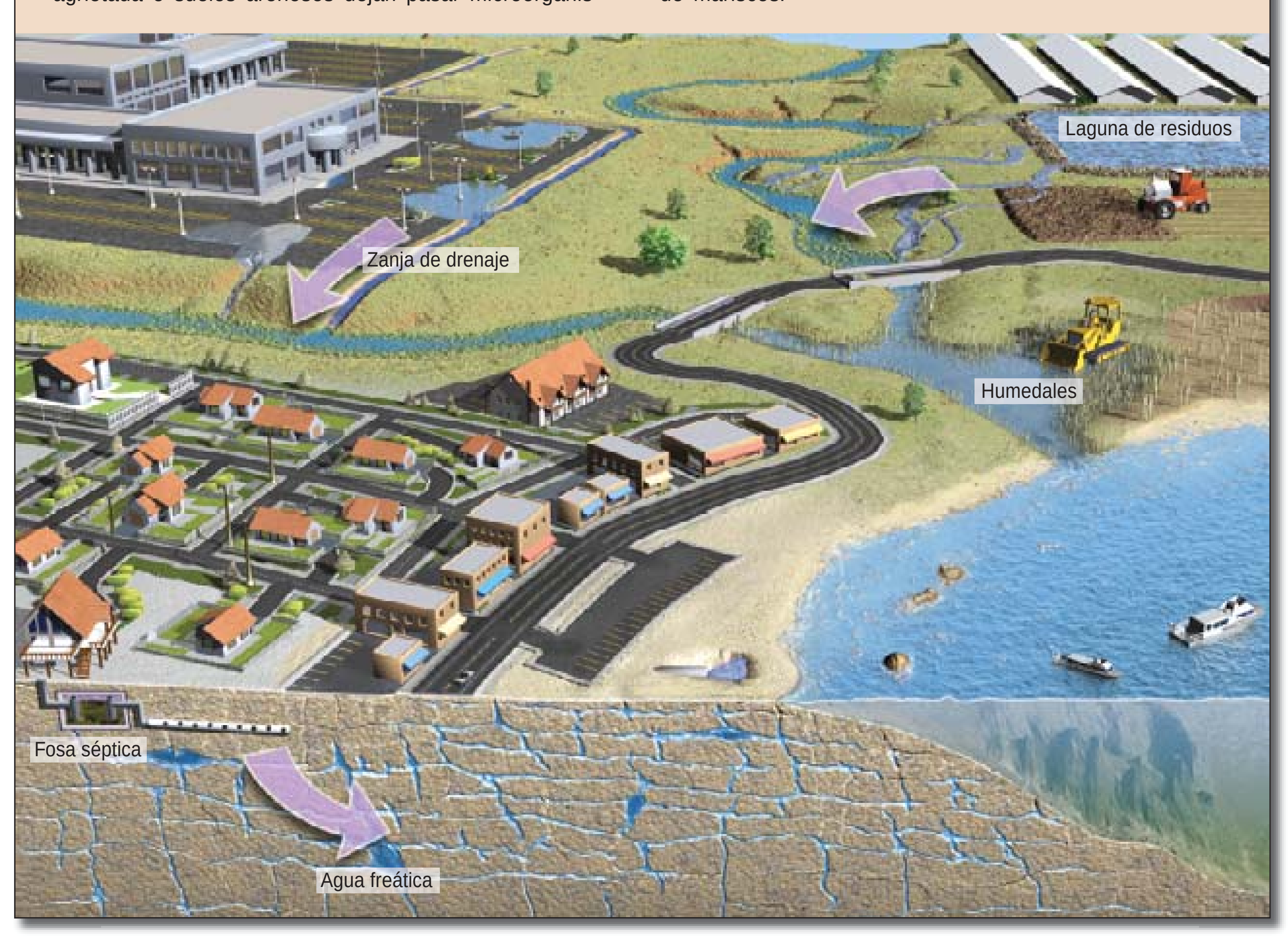
En fecha más reciente, mi laboratorio, de la Universidad de Carolina del Norte en Wilmington, analizó los datos procedentes de cinco condados costeros de aquel estado. Hallamos una clara correlación entre el crecimiento de la población humana y el cierre de viveros de moluscos. En 1984, cuando vivían 352.125 personas en esos cinco condados, se prohibió el marisqueo en 14.275 hectómetros cuadrados de agua. En 2003, la población total había ascendido a 501.596 personas; la superficie prohibida creció hasta 17.120 hectómetros cuadrados.

La contaminación bacteriana encierra, además, un grave riesgo para quienes practican la natación, el buceo, surf, esquí acuático, submarinismo, navegación y otras actividades recreativas. Si los coliformes fecales contaminan un lago, un río o una playa, todo el que se sumerja en esas aguas corre el riesgo de infectarse por microorganismos que penetren por la boca, la nariz, los ojos o heridas abiertas.

El contacto con el agua puede causar gastroenteritis, conjuntivitis (infecciones oculares), celulitis (irritaciones de la piel como la cercariosis cutánea o “prurito del nadador”), otitis, infec-

CIRCULACION DE MICROORGANISMOS FECALES

La mala planificación urbanística constituye la causa principal de la contaminación microbiana en las zonas costeras. La mayoría de los centros comerciales, por ejemplo, están rodeados por inmensas zonas de aparcamiento que conducen las aguas de escorrentía contaminadas a zanjas de drenaje. En numerosas urbanizaciones costeras, las fosas sépticas situadas sobre piedra caliza agrietada o suelos arenosos dejan pasar microorganismos fecales a las aguas freáticas. Las granjas de animales dispersan el estiércol en el campo o lo almacenan en lagunas; las lluvias intensas las arrastran hasta los arroyos cercanos. A menudo, nuevas construcciones destruyen humedales que, en condiciones normales, filtran los microorganismos patógenos que contaminan el agua antes de que éstos lleguen a las playas y a los viveros de mariscos.



DAVID FIERSTEIN

ciones respiratorias y afecciones de mayor gravedad como la hepatitis y el síndrome de Guillain-Barré (inflamación de los nervios periféricos que provoca parálisis). Entre las bacterias que causan esos trastornos figuran: *Escherichia coli*, *Clostridium perfringens* y diversas especies de *Enterococcus*, *Aeromonas*, *Campylobacter*, *Salmonella*, *Shigella* y *Yersinia*. Las aguas también transportan numerosos virus patógenos como los de la hepatitis A y Norwalk, además de *Cryptosporidium*, *Entamoeba*, *Giardia* y otros protozoos patógenos.

Agua no potable

En años recientes, varios de estos microorganismos han desencadenado graves crisis sanitarias en EE.UU. y Canadá. En 1993, la red de agua potable de Milwaukee se contaminó con *Cryptosporidium*; como resultado se produjo un brote con 87 muertos y 400.000 afectados. En 1999, un brote de *Escherichia coli* y *Campylobacter* causó dos muertes y afectó a 116 excursionistas del condado de Washington en Nueva York, que bebieron agua contaminada por los vertidos de una granja. En 2000, los residentes de Walkerton, en Ontario, fueron víctimas de infecciones por *E. coli* y *Campylobacter*: seis personas fallecieron y 2300 cayeron enfermas, sobre todo ancianos y niños. Una vez más, los microorganismos procedentes de los vertidos de explotaciones ganaderas habían pasado a la red urbana de agua potable.

Para evaluar el peligro de los agentes patógenos que transporta el agua, sobre todo en las playas, se mide la concentración de bacterias indicadoras de contaminación fecal. Cuando la concentración bacteriana supera los estándares de calidad establecidos, las autoridades sanitarias avisan de que el agua está contaminada o cierran las playas para evitar el baño y otras actividades recreativas. La EPA recomienda el uso de las bacterias *Enterococcus* como indicador de salubridad del agua costera y oceánica.

Según los criterios de la EPA, se considera que el agua de mar es in-

EL ESTADO DE NUESTRAS PLAYAS

España, al igual que los demás países de la Unión Europea, tiene la obligación de analizar quincenalmente, durante la temporada balnearia, la calidad de las aguas de baño. Para ello se determinan los parámetros fisicoquímicos y microbiológicos (microorganismos coliformes fecales y totales) que establece la Directiva Europea (76/160/CEE). La tabla muestra la clasificación de las zonas de baño españolas y de los países de la UE durante 2006 con respecto al cumplimiento de la Directiva.

***ZONAS DE BAÑO COSTERAS ESPAÑOLAS QUE INCUMPLIERON LA DIRECTIVA EUROPEA EN 2006**

Playa de la Charca	Andalucía
Playa del Cable	Andalucía
Playa de Guadarranque	Andalucía
Playa de Poniente	Andalucía
Playa de Les Huelgues	Asturias
Playa de la Concha	Cantabria
Playa de Oribarzar	País Vasco
Playa de Saturrarán	País Vasco
Playa de Toña	País Vasco
Playa de San Antonio	Islas Baleares
Playa de Pollença-1	Islas Baleares
Playa de Pollença-2	Islas Baleares
Cala Padera	Islas Baleares

	ESPAÑA		EUROPA	
	Zonas costeras	Zonas interiores	Zonas costeras	Zonas interiores
Incumplimiento[1]	13* (0,7 %)	8 (4,7 %)	217 (1,5 %)	284 (4,2 %)
Prohibición del baño	0	3 (1,8 %)	279 (1,9 %)	338 (5,1 %)
Muestreo insuficiente	7 (0,4 %)	5 (2,9 %)	57 (0,5 %)	145 (2,1 %)
Cumplimiento[2]	1843 (98,9 %)	155 (90,6 %)	13.792 (96,1 %)	5982 (88,6 %)
Excelentes[3]	1727 (92,7 %)	70 (40,9 %)	12.680 (88,4 %)	4305 (63,8 %)

[1]Superación de los valores imperativos de 10.000 UFC/100 ml (unidades que forman colonias por 100 mililitros de agua) de coliformes totales y 2000 UFC/100 ml de coliformes fecales en más del 5 % de las muestras.

[2]Cumplimiento con los valores imperativos de 10.000 UFC/100 ml de coliformes totales y 2000 UFC/100 ml de coliformes fecales en el 95 % de las muestras.

[3]Zonas de baño que cumplen los valores guía más restrictivos de la Directiva: 500 UFC/100 ml de coliformes totales y 100 UFC/100 ml de coliformes fecales en el 80 % de las muestras.





La playa mallorquina de Pollença (*izquierda*) ofrece una clara muestra de la sobreexplotación urbanística que sufre el litoral español. Por fortuna, quedan todavía playas exentas de grandes construcciones como la del Castell (*derecha*), en Palamós.

salubre cuando la media geométrica de cinco muestras recogidas durante 30 días supera los 35 UFC/100 ml de *Enterococcus*, o si cualquier muestra individual supera las 104 UFC/100 ml. En agua dulce, los límites de las concentraciones medias fijados por la EPA (estándares) son de 33 UFC/100 ml para *Enterococcus* y 126 UFC/100 ml para *E. coli*. Estas normas, sin embargo, quizá sean excesivamente laxas. La EPA estima que bañarse en una playa con la contaminación microbiana máxima aceptable entraña un riesgo de contraer enfermedades del 2 por ciento para los bañistas.

Aguas arriba

Desde finales de los años ochenta se estudian los daños ambientales que causa la impermeabilización de superficies. Nuestro laboratorio fue el primero en examinar los efectos en la concentración de bacterias fecales. Nos centramos en el condado de New Hanover, una zona de Carolina del Norte en rápido crecimiento. De 1990 a 2000, la población aumentó en un 25 por ciento; se espera que para 2020 haya crecido en otro 31 por ciento.

Durante el decenio pasado, estudiamos la calidad del agua en seis arroyos creados por la pleamar en urbanizaciones del condado. Analizamos más de 1000 muestras, en busca de correlaciones entre la concentración de bacterias (coliformes fecales

NUEVA DIRECTIVA EUROPEA, MAS EXIGENTE, PARA 2008

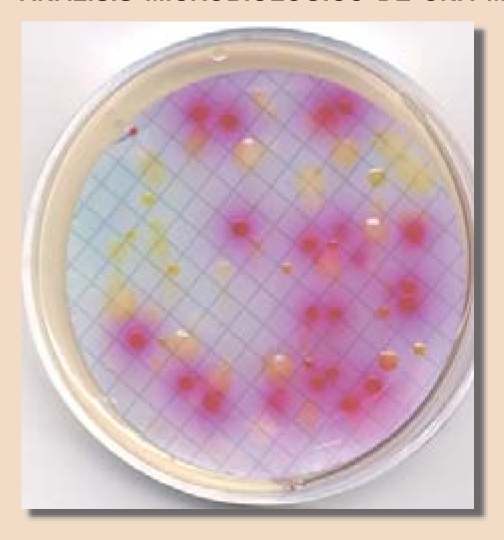
La Directiva Europea sobre la calidad de las aguas de baño (76/160/CEE) promulgada hace más de 30 años será reemplazada por la nueva Directiva sobre la gestión de las aguas de baño (2006/7/CE). Esta, más exigente, deberá aplicarse en toda Europa a partir de 2008. Entre otros cambios, exige la investigación inmediata de cualquier episodio de contaminación fecal, así como una pronta campaña informativa que alerte de estos episodios a los bañistas. Otra de las novedades de la nueva normativa son los indicadores bacterianos de contaminación fecal, que se adaptan a los conocimientos científicos actuales. Así, los microorganismos coliformes totales se suprimen, los coliformes fecales se reemplazan por *E. coli* y los estreptococos fecales se adaptan a la terminología actual de "enterococos intestinales". Los nuevos indicadores (*E. coli* y enterococos intestinales) coinciden con los que se utilizan hoy en Estados Unidos.

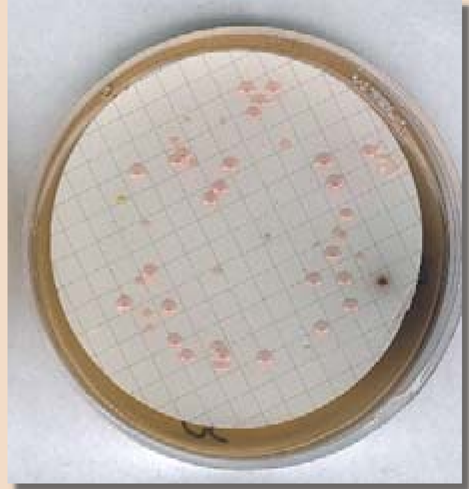
Los estándares de calidad microbiológica (concentración en unidades que forman colonias por 100 mililitros de agua, UFC/100 ml) establecidos para la clasificación de las zonas de baño en cuatro categorías (excelente, buena, suficiente e insuficiente) se basan en el cálculo del percentil 95 o 90 de los datos recopilados durante las tres o cuatro últimas estaciones balnearias. Estos niveles de referencia se inspiran en los valores que recomienda la Organización Mundial de la Salud, basándose en los escasos datos epidemiológicos existentes.

La falta de estudios epidemiológicos en el Mediterráneo y en aguas interiores de los nuevos países miembros ha motivado a la UE a financiar, dentro del VI Programa Marco, el proyecto de investigación epidemiológica EPIBATHE, que lidera la Universidad de Gales; la responsabilidad de la ejecución del estudio en España corre a cargo de Maria José Figueras, de la Universidad Rovira i Virgili. Los resultados de ese proyecto previstos para finales de 2008 permitirán revisar los estándares de calidad de la nueva directiva si se demuestra que los establecidos no protegen suficientemente la salud de los bañistas.

—Maria José Figueras,
Universidad Rovira i Virgili

ANALISIS MICROBIOLOGICO DE UNA MUESTRA DE AGUA DE BAÑO COSTERA





Estas placas muestran la presencia de colonias rojas de *E. coli* (*izquierda*) y de una colonia marrón oscuro de enterococos intestinales (*derecha*). Los análisis se realizaron con la metodología descrita en las normas ISO 9308-1 y ISO 7999-1, respectivamente, incluidas en la nueva Directiva (2006/7/CE).

MINISTERIO DE MEDIO AMBIENTE (playas). M.ª JOSE FIGUERAS (*cultivos*)

y *E. coli*) y los diversos parámetros demográficos y topológicos de los cauces correspondientes.

Encontramos que la cantidad (promediada) de coliformes era en general superior en los arroyos más poblados y en cuya cuenca había mayor proporción de suelo urbanizado. Pero la concentración de bacterias guardaba una correlación más estrecha con la presencia de superficies impermeables. En Futch Creek, con sólo el 7 por ciento del terreno circundante impermeabilizado, la concentración de coliformes era de 12 UFC/100 ml; en Bradley Creek, donde el asfalto y el hormigón cubrían el 22 por ciento de la vertiente, la concentración era más de siete veces mayor.

La correlación entre la abundancia de *E. coli* y el porcentaje de superficie impermeabilizada era también muy marcada. No fuimos los únicos en descubrirlo: A. Fred Holland, Dense M. Sanger y sus colaboradores, del Departamento de Recursos Naturales de Carolina del Sur, señalaron una correlación notable entre la concentración de coliformes y la impermeabilización de superficies en las vertientes de 22 arroyos de pleamar en el área metropolitana de Charleston.

Los resultados sugieren que el agua de escorrentía torrencial procedente de zonas urbanizadas ejerce un efecto multiplicador sobre la concentración bacteriana aguas

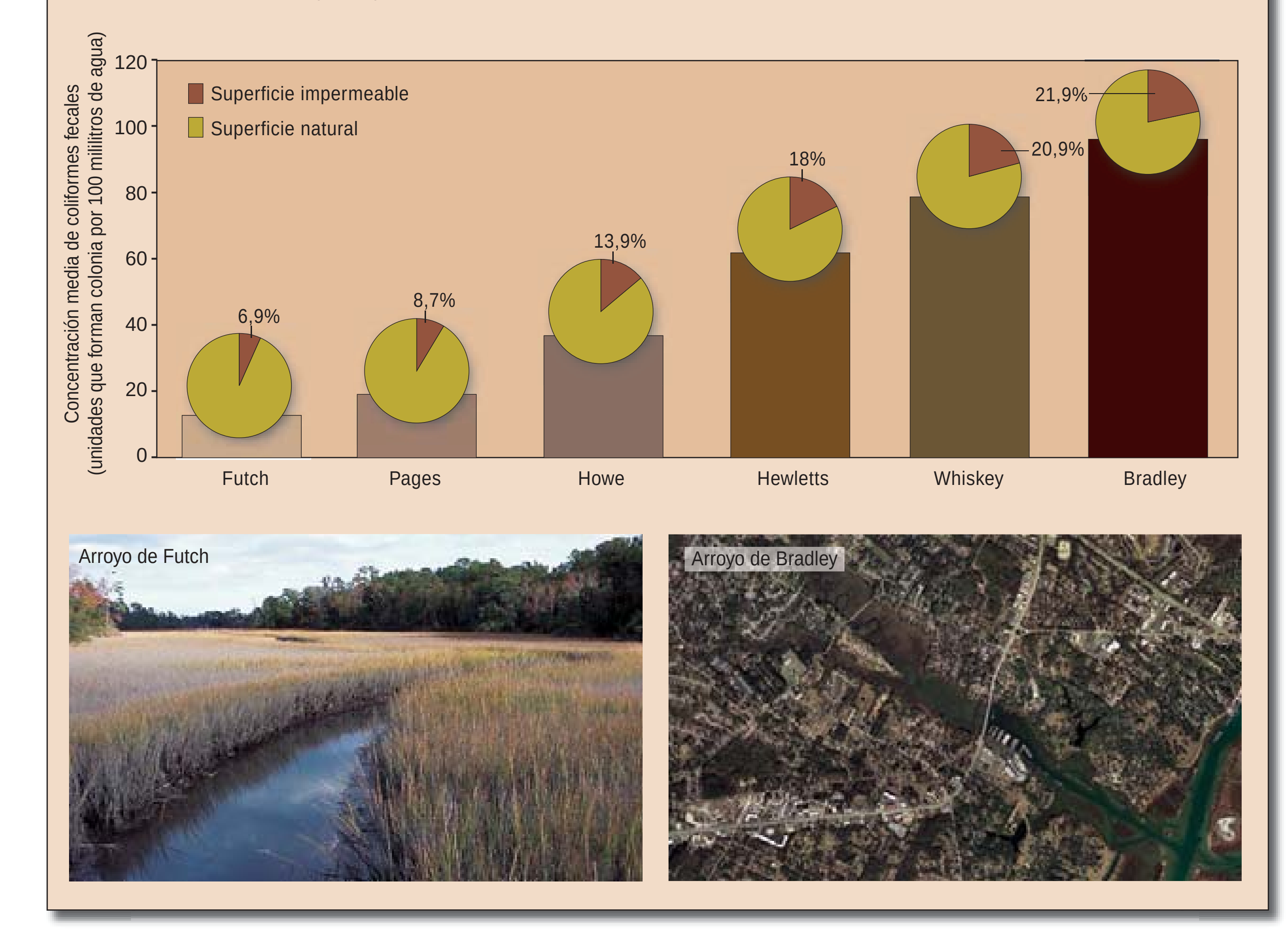
abajo. El caudal excesivo que proviene de grandes aparcamientos o zonas residenciales puede erosionar las zanjas de drenaje y las orillas de los cauces; ello aporta al agua sedimentos, que provienen también de solares edificables, donde el suelo se ha despojado de vegetación. Todos esos materiales en suspensión y otras partículas aumentan la turbidez de las aguas receptoras. Es más, los sedimentos, sobre todo las arcillas, se adhieren física y químicamente a contaminantes como el amonio, fosfato, metales, virus y bacterias fecales.

Las bacterias adheridas al suelo quedan protegidas de la radiación ultravioleta, letal para ellas. Además, obtienen del suelo carbono, nitrógeno, fósforo y otros nutrientes. Logran así recorrer largas distancias aguas abajo, unidas a los sedimentos. El estudio de los arroyos de pleamar en el Condado de New Hanover reveló una estrecha correlación entre la turbidez y la abundancia de coliformes fecales. En la bahía de Chesapeake, al oeste de Florida, la costa mediterránea y Australia se han obtenido resultados similares.

También los sedimentos que yacen en el fondo de aguas costeras someras constituyen un depósito de bacterias fecales y otros microorganismos. El equipo que dirige Lawrence B. Cahoon, colaborador mío en la Universidad de Carolina del Norte en Wilmington, ha hallado concentraciones elevadas de microorganismos patógenos (coliformes fecales, enterococos y estreptococos) en los sedimentos de arroyos de pleamar. Los microorganismos sobreviven largos períodos en los sedimentos porque éstos les ofrecen alimento y protección frente a la radiación ultravioleta. Además, dada la escasa profundidad de esos arroyos, la agitación de los sedimentos del fondo provoca la resuspensión de bacterias, de modo que se sobrepasan los estándares de seguridad establecidos para el contacto humano. Si sopla viento o hay niños y animales domésticos chapoteando, será fácil que se contaminen las aguas, sólo por el efecto de remover el barro del fondo. Hemos encontrado, además, concentraciones elevadas de microorganismos fecales en los sedimentos próximos a embarcaderos públicos, muy abundantes y

LOS EFECTOS DEL HORMIGON Y EL ASFALTO

Un estudio de seis arroyos de pleamar en el condado de New Hanover, en Carolina del Norte, reveló una elevada correlación entre la contaminación microbiana y el predominio de superficies impermeabilizadas como las de zonas de estacionamiento, carreteras y aceras. En Futch Creek (*izquierda*), las superficies impermeables no llegaban al 7 por ciento de la superficie de la cuenca; la concentración media de coliformes era allí muy inferior a la de Bradley Creek (*derecha*), donde el hormigón y el asfalto cubrían más de la quinta parte del terreno circundante.



LUCY READING-IKKANDA; MICHAEL A. MALLIN (*izquierda*); DEPARTAMENTO DE PLANIFICACION DEL CONDADO DE NEW HANOVER (*derecha*)

utilizados en las regiones costeras del sudeste.

Tratamiento de aguas residuales

Los sistemas de saneamiento mal diseñados de las regiones costeras contribuyen asimismo a la contaminación microbiana. Cuando los desagües pluviales descargan en las alcantarillas, éstas se desbordan por las lluvias intensas y vierten aguas residuales humanas sin depurar a ríos, lagos y bahías. Numerosos ayuntamientos abordan ahora ese problema mediante la construcción de sistemas separativos (separación en origen de los sistemas de alcantarillado y de drenaje pluvial). Emerge, no obstante, un nuevo dilema en zonas costeras donde, en ausencia de redes de alcantarillado, se eliminan las aguas residuales mediante fosas sépticas.

Ese es el caso de los Cayos de Florida (Florida Keys), donde operan más de 25.000 fosas sépticas. Predominan en ese terreno las formaciones kársticas: suelos de piedra caliza con abundantes grietas y cavidades producidas por la erosión; debido a su elevada porosidad, no filtran los vertidos con un alto contenido en bacterias de las fosas sépticas. Un equipo dirigido por John H. Paul y Joan B. Rose, de la Universidad de Florida del Sur, constató que los microorganismos fecales procedentes de las fosas sépticas en los Cayos atraviesan el suelo con facilidad; llegan a las aguas costeras en cuestión de horas.

El problema no es exclusivo de los Cayos. También los suelos arenosos presentan una porosidad elevada; cuando están saturados de agua, las bacterias y los virus se mueven holgadamente a través de los intersticios que separan los granos de arena.

Las zonas de suelo arenoso y alto nivel freático no son, pues, las más adecuadas para la construcción de fosas sépticas; sin embargo, debido a una mala planificación, se han instalado dichos sistemas de saneamiento en zonas costeras de rápido crecimiento, incluso en varias islas de la barrera arenosa que guarnece las costas del Atlántico y del Golfo. En un estudio de la calidad del agua realizado por Cahoon en el Condado de Brunswick (Carolina del Norte), en un terreno arenoso donde se cuentan hasta 20 fosas sépticas por hectárea, se detectaron niveles elevados de coliformes fecales en las aguas dulces y marinas alejadas de los núcleos residenciales densamente poblados. Y lo que es peor, los sistemas de encauzamiento y drenaje que atendían esas zonas urbanizadas parecía que facilitaban la descarga de bacterias fecales en las aguas próximas, incluidas las zonas de cultivo de moluscos.

Se han identificado otros patrones de contaminación microbiana. Erin K. Lipp, hoy en la Universidad de Georgia, y sus compañeros hallaron que, en poblaciones costeras del golfo de Florida como Charlotte Harbor y Sarasota Bay, la concentración de bacterias fecales en bahías y corrientes tributarias aumentaba de forma brusca con la bajamar. Los microorganismos se mueven con facilidad por los suelos arenosos y saturados que rodean las fosas sépticas, y fluyen hacia el mar a través de los arroyos cercanos.

A ello contribuyen no sólo las mareas, sino también las condiciones meteorológicas. Se descubrió que en los años húmedos correspondientes al fenómeno del Niño, el agua de la bahía de Tampa presentaba una contaminación por virus y bacterias fecales muy superior a la de los años secos. Ese efecto corresponde, una vez más, al resultado de mayores escorrentías y movimientos bajo la superficie a través de suelos saturados que rodean fosas sépticas mal emplazadas.

En el sur de California, donde se vierten al océano cantidades ingentes de aguas residuales urbanas, el grupo encabezado por Rachel Noble, de la Universidad de Carolina del Norte en Chapel Hill, comprobó que el número de zonas costeras que incumplían los estándares de calidad era diez veces mayor tras las lluvias que durante los períodos de sequía. A escala nacional, numerosos viveros de moluscos se cierran de forma automática durante algunos días o semanas después de haber llovido, pues se trata de zonas expuestas a la contaminación bacteriana que llega con las escorrentías torrenciales.

Pero no siempre cabe achacar la contaminación microbiana a las escorrentías urbanas y las filtraciones de fosas sépticas. Los arroyos que vierten al mar aguas procedentes de zonas rurales tienen también su parte de culpa, pues arrastran la contaminación fecal de origen animal que generan las explotaciones pecuarias. En la llanura costera que va desde Maryland hasta Florida, así como en ciertos lugares de la costa del Golfo, gran parte de las granjas tradicionales se han sustituido por explotaciones ganaderas intensivas, donde se crían cerdos, pollos y vacas en granjas. Estas utilizan su enorme producción de estiércol para abonar, mediante aspersión de líquido o simple dispersión, los campos circundantes. Si tal aspersión o dispersión se produce poco antes, o durante, una lluvia, los microorganismos fecales del estiércol llegan a los arroyos cercanos a través de la escorrentía superficial.

Un futuro más limpio

Si se quiere proteger las aguas costeras, es imprescindible que la industria de la construcción abandone las prácticas destructivas (deforestación, drenaje de humedales y pavimentación de grandes extensiones de terreno) y adopten estrategias de crecimiento sostenible.

En la planificación urbanística de nuevos centros turísticos, centros comerciales, complejos de oficinas y urbanizaciones, debe reducirse el uso de superficies impermeables y aumentar los espacios con vegetación. En un lugar donde abunden las zonas verdes entre los terrenos pavimentados se reducirán las escorrentías; el filtrado a través del suelo eliminará buena parte de los contaminantes que circulan por el asfalto y el hormigón.

Debe potenciarse la preservación —y, a ser posible, la ampliación— de los humedales para conservar este sistema de filtración natural de las escorrentías torrenciales. En un estudio sobre once cursos de agua en la llanura costera de Carolina del Norte, descubrimos que, en las cuencas con una presencia abundante de humedales (mayor del 13,5 por ciento de la superficie), la concentración de coliformes fecales apenas aumentaba en los períodos lluviosos. Los resultados indican que la preservación y ampliación de los humedales constituyen una medida eficaz para evitar la entrada de contaminación microbiana y sedimentos en suspensión en las corrientes que descienden. Los

constructores podrían colaborar en tal objetivo mediante la reducción de los vertidos de sedimentos que se generan en las obras.

Los promotores inmobiliarios deberían hacer uso de las nuevas técnicas que permiten reducir la escorrentía torrencial e incluso aplicar un tratamiento local. Por ejemplo, las zonas de estacionamiento pueden pavimentarse con un hormigón poroso, una sustancia semipermeable que deja pasar el agua hacia el suelo, al propio tiempo que proporciona suficiente soporte estructural a los automóviles. Asimismo, existen nuevos sistemas de recogida que encauzan la escorrentía pluvial de los aparcamientos hacia filtros que purifican el agua mediante capas de minerales absorbentes y materiales orgánicos. (También puede reducirse el tamaño de las grandes zonas de estacionamiento, proyectadas en su mayoría para las compras de fin de semana, no para el tráfico cotidiano.) Esas técnicas se aplicarían a los proyectos nuevos y a los ya acabados. Si una comunidad desea revitalizar sus playas o salvar la industria marisquera local mediante la reducción de la contaminación que transportan las escorrentías pluviales, debe instalar filtros en todas las áreas de estacionamiento, crear zonas verdes intermedias a lo largo de arroyos y canales, y proceder a la recuperación de humedales.

La ineficiencia de las fosas sépticas en suelos arenosos y en zonas de topografía kárstica indica claramente que las regiones de la costa requieren sistemas más avanzados de tratamiento de las aguas residuales. Algunos grupos conservacionistas han señalado, no obstante, que la construcción de sistemas de alcantarillado centralizados conducirá a un desarrollo todavía más masificado y, por tanto, a una mayor contaminación de las aguas. Por tanto, las autoridades municipales de las poblaciones costeras deberán imponer restricciones a la impermeabilización del suelo en los nuevos proyectos de construcción de sistemas de depuración. El límite podría fijarse, por ejemplo, entre el 10 y el 15 por ciento del área total (sobre todo en las proximidades de viveros de moluscos).

Las costas estadounidenses ofrecen un entorno maravilloso para disfrutar de unas vacaciones, así como para el emplazamiento de viviendas y negocios. Pero sin una cuidadosa planificación urbanística y una voluntad política que proteja los recursos costeros, esas playas magníficas, bahías deslumbrantes y plácidos arroyos de pleamar se convertirán en simples depósitos de aguas residuales. Debemos asegurar que el crecimiento incontrolado no arruine las mismas cualidades que atrajeron a tanta gente a la costa.

El autor

Michael A. Mallin se dedica a la ecología acuática. Ha estudiado la contaminación en aguas dulces, estuarios y sistemas marinos costeros. Es profesor de investigación en el Centro de Ciencias del Mar en la Universidad de Carolina del Norte en Wilmington. Centra su investigación en el impacto que los distintos usos del terreno, la incorporación de nutrientes y las catástrofes meteorológicas ejercen sobre la calidad del agua.

Bibliografía complementaria

EFFECT OF HUMAN DEVELOPMENT ON BACTERIOLOGICAL WATER QUALITY IN COASTAL WATERSHEDS. Michael A. Mallin, Kathleen E. Williams, E. Cartier Esham y R. Patrick Lowe en *Ecological Applications*, vol. 10, págs. 1047-1056; 2000.

DEMOGRAPHIC, LANDSCAPE, AND METEOROLOGICAL FACTORS CONTROLLING THE MICROBIAL POLLUTION OF COASTAL WATERS. Michael A. Mallin, Scott H. Ensign, Matthew R. McIver, G. Christopher Shank y Patricia K. Fowler en *Hydrobiologia*, vol. 460, págs. 185-193; 2001.

LINKAGES BETWEEN TIDAL CREEK ECOSYSTEMS AND THE LANDSCAPE AND DEMOGRAPHIC ATTRIBUTES OF THEIR WATERSHEDS. A. F. Holland, D. M. Sanger, C. P. Gawle, S. B. Lerberg, M. S. Santiago, G. H. M. Riekerk, L. E. Zimmerman y G. I. Scott en *Journal of Experimental Marine Biology and Ecology*, vol. 298, págs. 151-178; 2004.

El ORIGEN *de la* VIDA



ROBERT SHAPIRO

La aparición repentina de una macromolécula autorreplicante como el ARN era extremadamente improbable. Los iniciadores de la vida habrían sido entramados de reacciones químicas impulsados por una fuente de energía

Los descubrimientos extraordinarios inspiran conclusiones extraordinarias. Así, James Watson explica que inmediatamente después de que él y Francis Crick desvelasen la estructura del ADN, Crick salió disparado hacia el pub Eagle para contar a todo el que estuviese dispuesto a escucharle que habían descubierto el secreto de la vida. La estructura del ácido nucleico, una elegante doble hélice, merecía tamaño entusiasmo. Sus dimensiones permitían el almacenamiento de información en un lenguaje de cuatro compuestos químicos, las bases nitrogenadas, que llevaban a cabo una función análoga a la que desempeñan las 28 letras del alfabeto.

La información se almacenaba en dos largas cadenas. Cada una especificaba el contenido de la otra. Tal organización sugería un mecanismo para la reproducción: primero, las dos hebras de la doble hélice de ADN se separaban; luego, nuevos nucleótidos (los ladrillos del ADN y que contienen las bases) se alineaban a lo largo de las hebras separadas y se unían. Se obtenían, así, dos dobles hélices, cada una réplica de la original.

La estructura de Watson y Crick desencadenó una avalancha de descubrimientos sobre el modo en que operan las células. Esos hallazgos estimularon también las especulaciones en torno al origen de la vida. H. J. Muller, premio Nobel, escribió que el material genético era "material vivo, el representante actual de la primera forma de vida"; Carl Sagan lo representó como un "gen primitivo desnudo, con vida independiente e inmerso en una disolución diluida de materia orgánica". (En ese contexto, "orgánico" hace referencia a moléculas que contienen átomos de carbono enlazados, tanto las que se encuentran presentes en las formas de vida como las que no desempeñan ninguna función biológica.)

Se han propuesto multitud de definiciones de la vida. El comentario de Muller estaría en consonancia con la que se ha denominado "la definición de la NASA": la vida corresponde a un sistema químico autosuficiente, capaz de experimentar una evolución de tipo darwinista.

En su libro *El gen egoísta*, Richard Dawkins profundizó sobre esta imagen de la primera entidad viva: "En algún momento, surgió por casualidad una molécula de particular interés. La denominaremos el *replicador*. Puede que no fuese la molécula de mayor tamaño o complejidad, pero poseía la extraordinaria capacidad de autorreplicarse". Cuando Dawkins escribió esas palabras hace 30 años, el ADN era el candidato más probable para desempeñar ese papel. Más tarde, los investigadores volvieron su atención hacia otras moléculas como posibles replicadores primordiales. Sin embargo, otros expertos y yo mismo pensamos que el modelo del origen de la vida basado en una estructura replicadora es, en su esencia, erróneo. Proponemos un modelo alternativo, a nuestro parecer mucho más convincente.

Cuando el ARN dominaba la Tierra

Las primeras objeciones a la teoría del "ADN primordial" no tardaron en aparecer. La replicación del ADN requiere la intervención de cierta serie de proteínas, macromoléculas que difieren químicamente del ADN. Ambos, proteínas y ADN, se obtienen mediante la concatenación de subunidades o monómeros; pero mientras el ADN se fabrica a partir de nucleótidos, las proteínas se construyen con aminoácidos. Las proteínas llevan a cabo numerosas funciones celulares. Las enzimas, la subclase de proteínas más famosa, operan como catalizadores: aumentan la velocidad de procesos químicos que, sin ellas, se desarrollarían de una forma exasperantemente lenta e inútil para la vida. Pero las proteínas se fabrican siguiendo las instrucciones codificadas en el ADN.

Llegamos así al viejo dilema evolutivo: ¿qué fue primero, el huevo o la gallina? El ADN contiene la receta para la síntesis de proteínas. Sin embargo, esa información no se puede recuperar o copiar sin la ayuda de proteínas. ¿Cuál fue la primera macromolécula: las proteínas (la gallina) o el ADN (el huevo)?

Una posible respuesta surgió cuando se centró la atención en un nuevo candidato: el ARN, molécula versátil que se construye, al igual que el ADN, mediante el ensamblaje de nucleótidos, aunque desempeña numerosas funciones celulares. Ciertos ARN transportan la información desde el ADN hasta los ribosomas, orgánulos integrados en buena medida por otros tipos de ARN y donde se lleva a cabo la síntesis de proteínas. Mientras desempeña sus múltiples funciones, el ARN adopta la forma de una doble hélice, que recuerda al ADN, o la de una hebra sencilla plegada, a la manera de la proteína.

A principios del decenio de los ochenta, se descubrieron las ribozimas, moléculas similares a las enzimas pero hechas de ARN. Surgió así una sencilla solución a la adivinanza del huevo y la gallina. La vida comenzaría con la aparición de la primera molécula de ARN capaz de autorreplicarse.

En un artículo clásico publicado en 1986, Walter Gilbert, premio Nobel, escribía en la revista *Nature*: "Se puede considerar un mundo de ARN, donde sólo hay moléculas de ARN que catalizan la síntesis de sí mismas... Por tanto, la primera etapa de la evolución se desarrolla mediante moléculas de ARN que llevan a cabo las actividades catalíticas necesarias para autoensamblarse a partir de una sopa de nucleótidos". Según

KEN EWARD *BioGrafx*

esa hipótesis, el primer ARN autorreplicante que surgió de la materia inerte desempeñaba las funciones que hoy cumplen el ARN, el ADN y las proteínas.

Varios datos adicionales respaldan la hipótesis de que, en la evolución de la vida, el ARN surgió antes que las proteínas y que el ADN. Así, los cofactores, moléculas de un tamaño restringido, intervienen en las reacciones catalizadas por enzimas. Con frecuencia, esos cofactores portan unido un nucleótido de ARN, sin función manifiesta. Tales estructuras se han considerado "fósiles moleculares", reliquias de una época en que el ARN, ausentes ADN y proteínas, dominaba el mundo bioquímico.

Sin embargo, esa y otras pruebas respaldan sólo la conclusión de que el ARN precedió al ADN y a las proteínas. No aportan, sin embargo, ninguna información sobre el origen de la vida, que puede haber precisado de etapas previas al mundo de ARN, en las que imperaban otras entidades vivas. Para mayor confusión, suele emplearse la expresión "mundo de ARN" para referirse a ambos conceptos. Nosotros restringiremos el uso de "ARN primordial" a la participación del ARN en el origen de la vida, para distinguirlo de la aseveración de que el ARN apareció antes que el ADN y las proteínas.

Sopa primitiva de moléculas

La hipótesis del "ARN primordial" se enfrenta a una cuestión crucial: ¿Cómo surgió ese primer ARN autorreplicante? Hay obstáculos imponentes que se levantan contra la tesis de Gilbert, según la cual el ARN se formó a partir de una sopa de nucleótidos inertes.

Los nucleótidos del ARN corresponden a moléculas orgánicas complejas. Cada uno consta de varias subunidades: un azúcar, un grupo fosfato y una de las cuatro posibles bases nitrogenadas. Por tanto, cada nucleótido de ARN contiene entre 9 y 10 átomos de carbono, numerosos átomos de oxígeno y nitrógeno, y el grupo fosfato, todos ellos enlazados según un patrón tridimensional preciso. Existen múltiples formas alternativas de establecer tales conexiones; podrían generarse, por tanto, miles de otros posibles nucleótidos. Sin embargo, todas esas opciones alternativas no aparecen representadas en el ARN. Además, esa cifra resulta ínfima al lado de los cientos de miles o millones de moléculas orgánicas que presentan un tamaño similar y no corresponden a nucleótidos.

La hipótesis de que, pese a todo, se habrían formado los nucleótidos adecuados se inspira en un experimento archifamoso. En 1953 Stanley L. Miller aplicó una descarga eléctrica a una mezcla de gases sencillos que, supuestamente, representaban la atmósfera terrestre primitiva. Observó entonces la formación de aminoácidos. También se han identificado aminoácidos entre los componentes del meteorito Murchinson, que cayó sobre Australia en 1969. La naturaleza parece haber sido generosa a la hora de suministrar "ladrillos" moleculares. A partir de la extrapolación de esos resultados, algunos autores han imaginado que *todos* los componentes de la vida se encontraban ya en los meteoritos y podrían formarse fácilmente mediante experimentos parecidos al de Miller. Pero ése no es el caso.

Los aminoácidos —incluidos los que se obtienen en experimentos parecidos a los de Miller— evidencian una complejidad menor que la de los nucleótidos. Constan de un grupo amino (un nitrógeno y dos hidrógenos) y un grupo ácido carboxílico (un carbono, dos oxígenos y un hidrógeno), unidos ambos al mismo átomo de carbono. El más sencillo de los 20 aminoácidos que se utilizan para la síntesis de las proteínas naturales contiene sólo dos átomos de carbono; otros diecisiete constan de seis átomos de carbono o menos.

Los aminoácidos y otras sustancias producidas en abundancia en el experimento de Miller contenían entre dos y tres átomos de carbono. En cambio, nunca se ha detectado la presencia de ningún nucleótido, ni entre los productos de los experimentos con descargas eléctricas ni en los estudios de meteoritos. Parece que la materia inerte muestra preferencia por la formación de moléculas compuestas de pocos átomos de carbono; no favorece, por tanto, la síntesis de los nucleótidos necesarios para nuestra forma de vida.

Para salvar ese escollo, los defensores de la hipótesis del ARN primordial han desarrollado un nuevo campo de investigación, la síntesis prebiótica. Se aprestan a demostrar que el ARN y sus componentes pueden prepararse en el laboratorio a través de una secuencia de reacciones cuidadosamente controladas, con las condiciones y los materiales de partida que ellos consideran pertinentes.

Pero la investigación sobre la síntesis prebiótica tampoco carece de defectos. Imaginemos el cuadro siguiente: un golfista que habiendo jugado una bola por un circuito de 18 hoyos, supone que, en su ausencia, la bola también podría desplazarse por el campo dejada a su arbitrio. Se demuestra así la posibilidad del suceso; sólo cabe suponer que alguna combinación de fuerzas naturales (terremotos, vientos, tornados e inundaciones) daría lugar al mismo resultado, siempre y cuando transcurriera el tiempo suficiente. No hay que violar ninguna ley física para que se produzca la formación espontánea de ARN. Pero la probabilidad en contra es inmensa.

Algunos químicos han sugerido que primero surgió una molécula autorreplicante más sencilla, parecida al ARN, que dominó la vida en el

Resumen/*El origen de la vida*

- Las teorías sobre el origen de la vida se dividen en dos grupos principales: las que se basan en un "replicador primordial", una macromolécula autorreplicante (así, el ARN) que se habría formado por casualidad, y las teorías que se fundan en un "metabolismo primigenio", un sistema de moléculas pequeñas que formarían un entramado de reacciones químicas impulsado por una fuente de energía y capaz de evolucionar.
- Los defensores del modelo del replicador primordial deben explicar cómo se pudo formar una molécula de tamaña complejidad antes de que arrancara el proceso evolutivo.
- Los partidarios del modelo del metabolismo primigenio deben demostrar que, cuando la Tierra era joven, pudieron establecerse entramados de reacciones químicas con capacidad para el crecimiento y la evolución.

mundo anterior al ARN. Es de suponer que ese primer compuesto tendría también la capacidad catalítica del ARN. Dado que hasta la fecha no se ha hallado ningún rastro de tal hipotético replicador y catalizador primordial, el ARN tuvo que haber asumido, en algún momento posterior a su aparición, todas las funciones del mismo.

Sin embargo, aun en el caso de que la naturaleza hubiese proporcionado una sopa primitiva con los materiales de construcción adecuados (nucleótidos o algún sustituto más sencillo), su ensamblaje espontáneo para la creación de un replicador resulta tan inverosímil, que deja pequeña la imposibilidad de la preparación del caldo de cultivo.

Supongamos que la sopa con los ladrillos moleculares se ensambló en condiciones que privilegiaran la formación de cadenas. Estarían en compañía abundante de unidades defectuosas; la inclusión de una de ellas en una cadena naciente arruinaría la capacidad de la misma para operar como un replicador. La pieza defectuosa más sencilla tendría, por ejemplo, un solo "brazo" disponible para conectarse con las otras piezas de construcción, en vez de los dos requeridos para asegurar el crecimiento de la cadena.

En teoría, una naturaleza indiferente combinaría las piezas al azar; generaría una variedad inmensa de cadenas cortas, en vez de la cadena de mayor longitud y uniformidad geométrica que se precisa para posibilitar las funciones catalíticas y de replicación. La probabilidad de que ese último suceso tenga éxito es tan pequeña, que el hecho de que ocurra, aunque sólo fuere una vez, en cualquier lugar del universo visible, se consideraría una suerte de milagro.

Vida a partir de moléculas pequeñas

Christian de Duve, premio Nobel, realizó un llamamiento al rechazo de teorías basadas en probabilidades tan inconmensurablemente pequeñas que sólo pueden considerarse "milagros", fenómenos que se alejan del ámbito de la investigación científica. Desde este enfoque, debe descartarse la participación del ADN, el ARN, las proteínas y otras macromoléculas complejas en el origen de la vida. A cambio, la naturaleza inerte nos ofrece un repertorio muy diverso de moléculas pequeñas con las que trabajar.

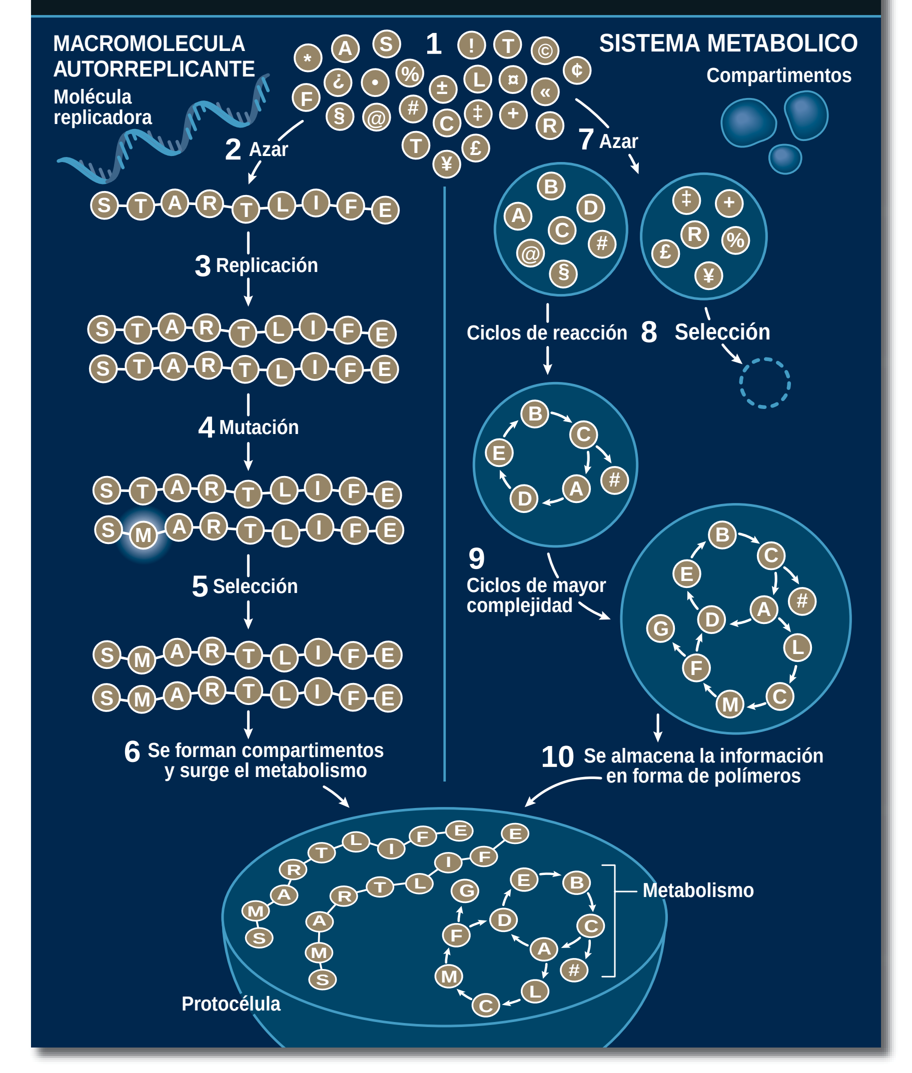
Por fortuna, durante decenios se han venido desarrollando otras teorías alternativas, basadas en la acción de esas moléculas. No parten de una definición genética de la vida, sino termodinámica, de acuerdo con un esquema propuesto por Sagan en la *Enciclopedia Británica*: una región circunscrita que aumenta el orden (disminuye la entropía) mediante ciclos impulsados por un flujo de ener-

DOS MODELOS SOBRE EL ORIGEN DE LA VIDA

A grandes rasgos, las teorías científicas sobre el origen de la vida se dividen en dos modelos enfrentados: el que supone la aparición de una macromolécula autorreplicante y el que se funda en un entramado metabólico de moléculas pequeñas. Ambas hipótesis parten de moléculas formadas en procesos químicos, no biológicos. En la figura, las moléculas aparecen representadas en forma de bolas marcadas con símbolos (*1*).

En el modelo del replicador, algunos de estos compuestos se unen entre sí en una cadena que, de un modo aleatorio, da lugar a una molécula (quizás algún tipo de ARN) autorreplicante (*2*). La macromolécula fabrica numerosas copias de sí misma (*3*); en ocasiones forma versiones mutantes, también autorreplicantes (*4*). Los replicadores mutantes que mejor se adaptan al medio suplantan a las versiones anteriores (*5*). En última instancia, ese proceso evolutivo debe conducir al desarrollo de compartimentos (como las células) y de un metabolismo, en el cual moléculas de menor tamaño utilizan energía para llevar a cabo procesos biológicos (*6*).

El modelo del metabolismo comienza con la formación espontánea de compartimentos (*7*); algunos contienen mezclas de los compuestos iniciales que establecen un entramado de reacciones químicas (*8*) que, con el tiempo, aumentan en complejidad (*9*). Por último, el sistema debe dar el salto que le permita almacenar en polímeros la información (*10*).



JEN CHRISTIANSEN

gía podría considerarse algo vivo. Ese enfoque basado en moléculas de tamaño reducido radica en las ideas del bioquímico soviético Alexander Oparin. Si bien difieren en ciertos detalles, las propuestas sobre el origen metabólico de la vida cumplen cinco principios básicos comunes, que presento a continuación (junto con algunas ideas de mi cosecha).

1. Entre lo vivo y lo inerte se necesita una barrera de separación. La vida se caracteriza por su elevado grado de organización, a pesar de que la segunda ley de la termodinámica requiere que el universo se mueva en una dirección en la que aumente el desorden, o entropía. Existe, sin embargo, una excepción a esa ley: la entropía puede disminuir en determinada región siempre que se produzca un mayor incremento de entropía en sus alrededores. Cuando las células crecen y se multiplican, convierten la energía química o la radiación en calor. El calor liberado aumenta la entropía del entorno, lo que compensa la disminución que tiene lugar en los sistemas vivos. La barrera mantiene esa división del mundo entre los compartimentos que contienen vida y el entorno inerte donde deben hallar sustento.

Ahora, las células se rodean de una membrana de doble capa lipídica que las separan de su entorno. Cuando comenzó la vida, es probable que algún rasgo de la naturaleza operase con el mismo propósito. David W. Deamer, de la Universidad de California en Santa Cruz, ha observado en meteoritos estructuras con aspecto de membrana, lo que parece respaldar tal supuesto. Otras hipótesis hablan de barreras naturales que no utilizan hoy los seres vivos: membranas de sulfuro de hierro, superficies rocosas (en donde las interacciones electrostáticas apartan de su entorno moléculas seleccionadas), pequeños estanques y aerosoles.

2. Una fuente de energía debe impulsar el proceso de organización. Para mantenernos vivos consumimos hidratos de carbono y grasas, y los combinamos con el oxígeno que inhalamos. Los microorganismos son más versátiles: utilizan minerales, en vez de comida u oxígeno. En cualquier caso, las transformaciones que tienen lugar corresponden a reacciones redox (de "reducción" y "oxidación"). Implican la transferencia de electrones desde una sustancia rica en electrones (reducida) hacia una sustancia pobre en electrones (oxidada). Las plantas captan directamente la energía solar y la adaptan para las funciones biológicas.

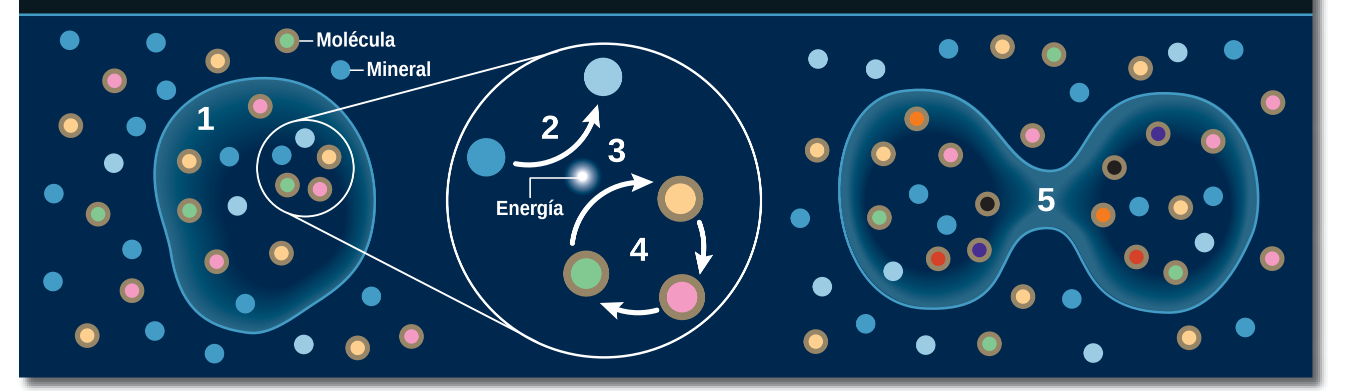
En circunstancias especiales, las células utilizan otras formas de energía: por ejemplo, las diferencias de acidez a ambos lados de una membrana. Sin embargo, en otras partes del universo podrían utilizarse otras formas de energía para la vida como la radiactividad o los gradientes térmicos.

3. Un mecanismo de acoplamiento debe conectar la liberación de la energía con el proceso de organización que produce y sustenta la vida. La liberación de energía no da lugar, necesariamente, a un resultado útil. Durante la combustión de la gasolina en el interior de los cilindros de un automóvil, se libera energía química; pero el vehículo no se moverá a menos que se utilice esa energía para hacer girar las ruedas. Se necesita una conexión mecánica, o acoplamiento. Cada día, en el interior de nuestras células, degradamos kilos del nucleótido ATP. La energía liberada mediante esa reacción sirve para impulsar procesos necesarios para nuestra bioquímica que, sin tal auxilio, no tendrían lugar o serían demasiado lentos. La conexión se consigue cuando las reacciones comparten un intermediario común y el proceso se acelera mediante la intervención de una enzima. La hipótesis del "micrometabolismo" primigenio supone en la naturaleza la existencia de las reacciones acopladas y los catalizadores primitivos necesarios para iniciar la vida.

4. Debe establecerse un entramado de reacciones químicas que facilite la adaptación y la evolución. Llegamos ahora al quid de la cuestión. Imaginemos, por ejemplo, que una reacción redox energéticamente

CINCO REQUISITOS PARA UN ORIGEN METABOLICO DE LA VIDA

Para que las moléculas de parvo tamaño adquieran cierto tipo de vida —vale decir, la creación de más orden, en regiones circunscritas y mediante ciclos químicos impulsados por un flujo de energía—, deben darse cinco procesos, por lo menos. En primer lugar, ha de formarse una barrera que separe del entorno inerte la región viva (*1*). Luego, se exige una fuente de energía accesible; en este ejemplo, un mineral (*azul*) que experimenta una reacción exotérmica (*2*). Se precisa que la energía liberada impulse una reacción química (*3*). Se necesita, asimismo, la creación de un entramado de reacciones químicas cuya complejidad vaya en aumento para favorecer la adaptación y la evolución (*4*). Por fin, la red química debe incorporar material más rápidamente de lo que lo consume; sin olvidar que los compartimentos han de poder reproducirse (*5*). No se demanda ninguna molécula que almacene información (del estilo del ARN o el ADN); la herencia queda almacenada en la identidad y la concentración de los compuestos químicos del entramado.



JEN CHRISTIANSEN

favorable en un mineral impulsa, en el interior de un compartimento, la conversión de una sustancia orgánica A en otra, B. Una transformación clave, que denominaremos "reacción impulsora", pues actúa a modo de motor del proceso de organización. Si B se volviera a convertir en A o abandonara el compartimento, el proceso no conduciría a una mayor organización. Ahora bien, si una ruta química vuelve a convertir B en A (supongamos que B se convierte en C, C en D y D en A), las etapas de este ciclo se ven favorecidas para seguir operando porque restituyen el suministro de A, continuando así con la liberación útil de energía mediante la reacción mineral.

Pueden darse también reacciones ramificadas, como las que convierten la molécula D en otro compuesto químico E (o viceversa), en una ruta que no forma parte del ciclo ABCD. Puesto que el ciclo viene impulsado, se favorece la reacción de E a D, que incorpora material en el ciclo y maximiza la liberación de energía que acompaña a la reacción impulsora.

El ciclo se adaptaría también a circunstancias cambiantes. De niño, me fascinaba la forma en que el agua que salía de una boca de riego con fugas encontraba, cuesta abajo, una ruta hacia la alcantarilla más cercana. Si el camino se veía bloqueado por hojas caídas o algún otro obstáculo, el agua retrocedía hasta encontrar otra ruta que lo rodease. Del mismo modo, si un cambio en la acidez o en cualquier otra variable ambiental interrumpiese una de las etapas de la ruta que lleva de B a A, el material retrocedería hasta encontrar otro camino. Una serie de cambios de ese tipo convertiría el ciclo original en un entramado. Tal exploración del panorama químico basada en el ensayo y error promovería la aparición de compuestos que catalizarían las etapas cruciales del ciclo, con el consiguiente aumento del rendimiento energético de la red metabólica.

5. El entramado de reacciones químicas debe crecer y reproducirse. Para sobrevivir y crecer, el entramado metabólico debe captar materia más rápidamente de lo que la consume. La difusión de los elementos del entramado hacia el exterior del compartimento, que tendrá lugar en mayor o menor medida, se ve favorecida por la entropía. Algunas reacciones secundarias pueden producir gases, que se escapan, o formar alquitranes, que abandonarán la disolución. Si tales procesos combinados superasen la velocidad con la que el entramado incorpora materia, éste se extinguiría. El agotamiento del combustible externo tendría el mismo efecto.

Podemos imaginar, en la Tierra primitiva, una situación en la que se produjeran multitud de inicios del tipo expuesto, en los que intervendrían numerosas reacciones impulsoras alternativas y muchas fuentes externas de energía. Al final, se afianzaría la que mostrara mayor resistencia y autonomía.

En última instancia, debe desarrollarse un sistema de reproducción. Si nuestro entramado metabólico está delimitado por una membrana lipídica, podría escindirse mediante fuerzas físicas, tras haber alcanzado un tamaño suficiente. (Freeman Dyson, del Instituto de Estudios Avanzados de Princeton en New Jersey, ha descrito ese sistema como "el mundo en una bolsa de basura", en contraste con el "escenario impoluto y hermoso" del "mundo de ARN".) Un sistema que opera en un compartimento en el interior de una roca puede desbordarse hacia compartimentos adyacentes. Cualquiera que sea el mecanismo, esa dispersión en unidades separadas protege al sistema de su extinción total por un suceso destructivo local. Una vez establecidas las unidades independientes, seguían diferentes trayectorias evolutivas y competirían entre sí por las materias primas; se habría producido la transición de una vida surgida a partir de la materia inerte, merced a la acción de una fuente de energía, a una vida que se adapta al entorno mediante evolución darwinista.

Cambio de paradigma

Los sistemas de esa clase corresponden al modelo que atribuye el origen de la vida a un "metabolismo primordial"; carecen, por tanto, de un mecanismo para la herencia. En otras palabras, no cuentan con una

Preguntas frecuentes

¿Implica la hipótesis de un metabolismo primigenio un origen de la vida único o múltiples orígenes independientes?

Según el modelo basado en un metabolismo primordial, parece más viable un escenario con múltiples orígenes. En el libro "Life beyond Earth" que *Scientific American* publicó en 1980, Gerald Feinberg y el autor del artículo discuten la posibilidad de vida alienígena (que no está basada en el ARN, ADN y otros elementos bioquímicos con los que estamos familiarizados). Los expertos que asistieron al congreso que organizó Paul Davies en la Universidad estatal de Arizona en diciembre de 2006 concluyeron que es posible que la vida alienígena exista incluso en nuestro planeta, aunque no la hayamos detectado. La inmensa mayoría de los microorganismos que se observan con un microscopio no crecen en los medios de cultivo al uso; por tanto, siguen sin estar caracterizados. Los microorganismos alienígenas quizás existan también en hábitats de la Tierra que resultan demasiado extremos para las formas de vida conocidas, incluso las más resistentes.

¿Por qué es necesaria la demostración experimental de la hipótesis del metabolismo primordial? ¿No bastan las simulaciones por ordenador?

Stuart Kauffman, Doron Lancet y otros investigadores han utilizado simulaciones por ordenador para mostrar la viabilidad de los ciclos de reacción autosuficientes. Pero esas simulaciones no especifican la concentración de los reactivos, ni las condiciones de reacción necesarias para la aparición de redes químicas autosuficientes. No conocemos todavía todas las rutas de reacción que pueden seguir las mezclas de compuestos orgánicos sencillos y, mucho menos, sus constantes de reacción termodinámicas. Pero aunque contáramos con tales datos, no bastaría una simulación teórica. La última palabra la tiene siempre la demostración experimental.

¿Por qué todas las biomoléculas muestran preferencia por una de sus formas quirales?

Si la transformación mineral que impulsa el ciclo de reacciones que aparece en el artículo seleccionase una de las imágenes especulares del compuesto químico A, el producto B y otros miembros del ciclo también existirían sólo en una de las formas especulares. El control de la quiralidad resulta crucial cuando hay moléculas quirales pequeñas que se conectan entre sí para crear moléculas de mayor tamaño. Una enzima moderna puede estar formada por la unión de 100 aminoácidos, todos con la misma quiralidad (L-aminoácidos). Si en una región sensible de la enzima hubiese un D-aminoácido, en vez de su imagen especular (la forma L), cambiaría la forma de la enzima y podría perder su función.

molécula o estructura encargada de duplicar y transmitir a sus descendientes la información que contienen (la herencia). Sin embargo, un conjunto de elementos de tamaño reducido contiene la misma información que una lista que los enumere. Pensemos en la bolsa de bienes comprados en el supermercado: contiene la misma información que la lista de la compra. Doron Lancet, del Instituto Weizmann de Ciencias en Rehovot, ha denominado "genoma composicional" a la herencia almacenada en forma de moléculas pequeñas, en vez de en una lista como el ADN o el ARN.

La hipótesis que basa el origen de la vida en un sistema de moléculas pequeñas impone ciertas exigencias a la naturaleza (un compartimento, un suministro externo de energía, una reacción impulsora acoplada a ese suministro, un metabolismo que incluya esa reacción y un mecanismo sencillo de reproducción). Unos requerimientos que son comunes en la naturaleza. E inmensamente más probables que las complejas y múltiples etapas que requiere la formación de una molécula autorreplicante.

A lo largo de los años, se han presentado numerosos artículos teóricos sobre modelos basados en un "metabolismo primordial", pero pocos trabajos experimentales que los respalden. Además, esos escasos estudios experimentales suelen demostrar la viabilidad sólo de etapas particulares dentro del ciclo propuesto. La mayoría de los datos proceden de los trabajos de Günter Wächtershäuser y sus colaboradores, de la Universidad Técnica de Múnich. Han demostrado fases de un ciclo que implica la combinación y separación de aminoácidos en presencia de sulfuros metálicos que operan a modo de catalizadores. La energía que impulsa las transformaciones se suministra mediante la oxidación del monóxido de carbono en dióxido de carbono. Pero no se ha demostrado todavía la viabilidad de un ciclo completo ni su capacidad para ser autosuficiente y experimentar una evolución ulterior. Para confirmar la validez de la hipótesis del micrometabolismo primigenio hacen falta experimentos en los que esas tres características queden patentes sin ambigüedad.

La tarea principal consiste en la identificación de las posibles reacciones impulsoras: transformaciones de moléculas pequeñas (A en B, en el ejemplo anterior) acopladas a una fuente externa de energía abundante (como la oxidación del monóxido de carbono o de un mineral). Una vez identificada una posible reacción impulsora, no habría necesidad de especificar por adelantado el resto del sistema. Los componentes seleccionados (incluida la fuente energética), junto con una mezcla de otras moléculas pequeñas producidas mediante procesos naturales (y con una probabilidad elevada de haber existido en abundancia en la Tierra primitiva), se combinarían en un balón de reacción.

Si se estableciese un entramado de reacciones químicas en evolución, cabría esperar que la concentración de los reactivos aumentaría y se alteraría con el tiempo. Podrían aparecer nuevos catalizadores que acelerasen la velocidad de las reacciones clave, al tiempo que mermara la cantidad de compuestos irrelevantes. El reactor necesitaría un dispositivo de entrada (para facilitar el aflujo de energía y materias primas) y una salida (para la expulsión de los productos de desecho y de las sustancias químicas que no formaran parte del entramado).

En los experimentos de ese tenor, los fallos se descubren en seguida. Podría acontecer que la energía se disipara sin que se produjeran cambios significativos en la concentración de otros compuestos químicos, o que las sustancias se convirtieran en un alquitrán, que obturaría el aparato. El éxito, en cambio, vendría a demostrarnos los primeros pasos en el camino que conduce a la vida. No hay razón para que esas etapas iniciales fueran exactamente las mismas que tuvieron lugar en la Tierra primitiva. Reviste interés mayor la demostra-

En defensa del modelo del ARN primigenio

A pesar de que algunos expertos sostienen que el modelo del ARN primordial es estéril, porque es altísimamente improbable que el ARN surgiese de forma espontánea, la investigación está recogiendo datos que respaldan esta hipótesis.

Reconozcamos de entrada que la mayoría de las sustancias orgánicas, cuando se les aporta energía (la de un rayo o el calor de un volcán), se convierten en algo parecido al asfalto, más apropiado para la construcción de carreteras que para el inicio de la vida. Pero los modelos que sugieren un origen "metabólico" de la vida, desde el momento en que se basan en *cualquier* tipo de moléculas reales, también se enfrentan a esa paradoja: las moléculas suficientemente reactivas como para intervenir en un metabolismo lo son, asimismo, para descomponerse. La solución no es sencilla.

Igual que otros, mi grupo de investigación ha regresado al principio científico según el cual la investigación del modo en que pudo haber surgido el ARN en la Tierra debe llevarse a cabo en el laboratorio.

El azúcar ribosa, la "R" del ARN, proporciona un buen ejemplo de cómo un problema declarado "irresoluble" puede ser en realidad un problema "todavía sin resolver". Durante muchos años resultó "imposible" fabricar ribosa mediante síntesis prebiótica (reacciones entre mezclas de moléculas que podían haber existido en una Tierra prebiótica) debido a la presencia en la molécula de un grupo carbonilo, un átomo de carbono unido mediante un doble enlace a un átomo de oxígeno. El grupo carbonilo confiere una reactividad "beneficiosa" (que le capacita para intervenir en el metabolismo), aunque también otra "perjudicial" (que le faculta para formar asfalto). Un decenio atrás, Stanley L. Miller llegó a la conclusión de que la inestabilidad de la ribosa debida al grupo carbonilo descartaba el uso de este y otros azúcares como agentes prebióticos. Tales compuestos, por tanto, no podían ser componentes del material genético primigenio.

Pero las sopas prebióticas necesitan recipientes fabricados con los minerales apropiados, no con el pyrex del material de laboratorio. Un recipiente atractivo lo encontramos en el Valle de la Muerte. En los primeros tiempos de su existencia, la zona, que alternaba estados húmedo y seco, era rica en moléculas orgánicas suministradas por el propio planeta y repleto de minerales que contenían boro. ¿Por qué resaltar el boro? Porque ese elemento estabiliza hidratos de carbono como la ribosa. Asimismo, si se mezcla borato (un óxido de boro) con los compuestos orgánicos que abundan en los meteoritos y se calienta mediante descargas eléctricas se forma abundante ribosa a partir de formaldehído; además, la ribosa no se descompone.

El hecho de que se haya encontrado una solución tan sencilla a un problema declarado "irresoluble" no significa que la primera forma de vida usara el ARN para hacer genética. Pero sí debería hacernos recapacitar cuando alguien sugiera el abandono de una línea de investigación sencillamente porque todavía no se han resuelto algunos de los aspectos problemáticos.

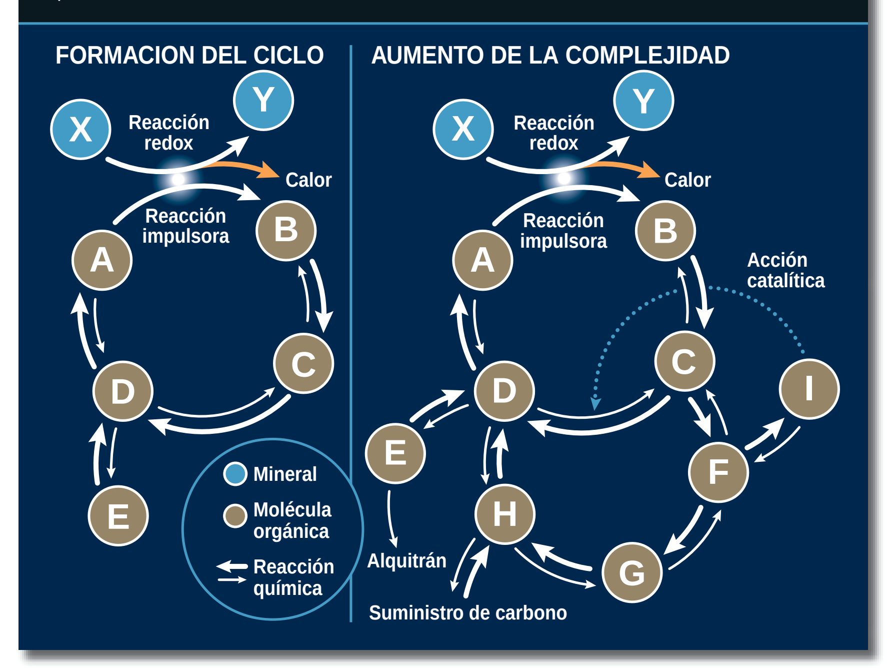
—Steven A. Benner, Instituto Westheimer para la Ciencia y la Tecnología en Gainesville, Florida

EVOLUCION DE REDES QUIMICAS

La hipótesis de un origen metabólico de la vida requiere el establecimiento de un entramado de reacciones químicas cuya complejidad aumente y que se adapte a los cambios del entorno.

FORMACION DEL CICLO: Una fuente de energía (en este caso, la reacción redox que convierte el mineral *X* en el mineral *Y*) se acopla a una reacción que transforma la molécula orgánica *A* en *B*. Reacciones ulteriores (*B* en *C*, *C* en *D*...) generan un ciclo que regresa a *A*. Las reacciones en las que participan especies moleculares ajenas al ciclo (*E*) tenderán a aportar más material al sistema.

AUMENTO DE LA COMPLEJIDAD: Si un cambio en las condiciones inhibe una de las reacciones del ciclo (por ejemplo, la transformación de *C* en *D*), se pueden explorar otras vías. En este caso, se ha dado un rodeo mediante el cual *C* se convierte en *D* a través de los intermediarios *F*, *G* y *H*. Otra solución consistiría en la incorporación de un catalizador (*I*) cuya actividad (*línea punteada*) desbloquee la transformación de *C* en *D*. Para sobrevivir, la red debe captar del entorno moléculas que contengan átomos de carbono y hacerlo con mayor prontitud de lo que las pierde por difusión o mediante reacciones secundarias, como la formación de alquitranes, que abandonan la disolución.



JEN CHRISTIANSEN

El autor

Robert Shapiro es profesor emérito de química e investigador en la Universidad de Nueva York. Centra su trabajo en el estudio de la química del ADN. En concreto, ha investigado el daño químico que provoca el medio sobre nuestro material hereditario, causando cambios mutagénicos y cancerígenos.

Bibliografía complementaria

PREBIOTIC CYTOSINE SYNTHESIS: A CRITICAL ANALYSIS. IMPLICATIONS FOR THE ORIGIN OF LIFE. Robert Shapiro en *Proceedings of the National Academy of Sciences USA*, vol. 96, págs. 4396-4401, abril 1999.

PREBIOTIC CHEMISTRY AND THE ORIGIN OF THE RNA WORLD. Leslie E. Orgel en *Critical Reviews in Biochemistry and Molecular Biology*, vol. 39, págs. 99-123, 2004.

EVOLUTIONARY POTENTIAL AND REQUIREMENTS FOR MINIMAL PROTOCELLS. E. Szathmáry, M. Santos y C. Fernando en *Topics in Current Chemistry*, vol. 259, págs. 167-211, 2005.

PROGRESS TOWARD UNDERSTANDING THE ORIGIN OF THE RNA WORLD. Gerald F. Joyce y Leslie E. Orgel en *The RNA World*, tercera edición, págs. 23–56. Dirigido por Raymond F. Gesteland, Thomas R. Cech y John F. Atkins. Cold Spring Harbor Laboratory Press, 2006.

SMALL MOLECULE INTERACTIONS WERE CENTRAL TO THE ORIGIN OF LIFE. Robert Shapiro en *Quarterly Review of Biology*, vol. 81, págs. 105-125, junio 2006.

ción del principio general y que éste pueda someterse a investigaciones ulteriores. Caben quizá numerosas vías potenciales que desemboquen en la vida. Correspondería al entorno local escoger la indicada.

El conocimiento de las fases iniciales del proceso que condujo a la aparición de vida no nos revelaría qué sucesos específicos condujeron hasta los organismos basados en ADN, ARN y proteínas.

Además, puesto que sabemos que la evolución no anticipa eventos futuros, cabe imaginar que los nucleótidos aparecieran en el metabolismo con una finalidad distinta, quizá como catalizadores o como contenedores para el almacenamiento de energía química (el nucleótido ATP sigue desarrollando esa función en la actualidad). Alguna circunstancia o suceso azaroso habría conducido a la unión de nucleótidos para formar ARN.

La función más obvia del ARN moderno es la de servir de elemento estructural que facilita la formación de enlaces entre aminoácidos durante la síntesis de proteínas. Los primeros ARN podrían haberse utilizado para el mismo fin, pero sin mostrar ninguna preferencia hacia aminoácidos específicos. Harían falta muchas más etapas en la evolución para "inventar" los elaborados mecanismos de replicación y síntesis de proteínas específicas que operan hoy en los seres vivos.

Si se confirmase el paradigma basado en el micrometabolismo, cambiarían nuestras expectativas sobre el lugar de la vida en el universo. Un inicio de la vida muy improbable, como el escenario del ARN primordial, implica que estamos solos en el universo. En palabras del bioquímico Jacques Monod, "el universo no estaba preñado de vida ni la biosfera de hombres. Nuestro número salió en el juego de la ruleta".

En cambio, la posibilidad alternativa que se funda en un metabolismo primordial está en armonía con la visión del biólogo Stuart Kauffman: "Si todo esto es verdad, la vida es mucho más probable de lo que hemos supuesto. No somos los únicos habitantes el universo; lo más probable es que lo estemos compartiendo con compañeros que no conocemos todavía".

Hacia una anestesia más segura

La investigación sobre la potencia y el riesgo de la anestesia general abre nuevas vías para el desarrollo de fármacos selectivos, más seguros y sin efectos secundarios

En un filme de suspense reciente, un hombre joven, sometido a una intervención a corazón abierto, se despierta, pero es incapaz de gritar o moverse. Una situación que se presta a lances efectistas de todo tipo. Con todo, la escena inicial no es pura ficción. Uno o dos de cada 1000 pacientes sufren episodios de consciencia intraoperativa bajo anestesia general. Suelen ser incidentes breves; en general, no implican dolor ni sufrimiento. Pero ponen de manifiesto una de las carencias de los fármacos anestésicos, incluidos los de la última generación. La anestesiología ha evolucionado en una suerte de arte, pues el conocimiento del modo en que operan los anestésicos y de cómo mejorarlos ha ido a la zaga de la mayoría de las demás áreas farmacológicas.

La mayor parte de los anestésicos modernos comparten propiedades estructurales y efectos clínicos con el éter, cuyas propiedades anestésicas fueron demostradas en 1846 por William Morton, dentista de Boston. Desde entonces, el uso de la anestesia general se ha expandido hasta los 40 millones de pacientes anuales, sólo en Estados Unidos. Sin embargo, los progresos en el cuidado anestésico desde los tiempos de William Morton se han debido, sobre todo, al desarrollo de las técnicas de administración de fármacos complejos y a las estrategias para el control del riesgo y los efectos secundarios de la anestesia.

Los anestésicos generales modernos son los más potentes depresores de la actividad del sistema nervioso que se emplean en medicina. Afectan incluso a la regulación respiratoria y a la función cardíaca. Su margen de seguridad es, por tanto, estrecho; corresponde a la diferencia entre la dosis terapéutica y una dosis tóxica, puede que letal. Esa es una de las razones por las que los individuos con una función pulmonar o cardiovascular inestable deben recibir una dosis de anestesia inferior a la normal: hablamos de personas que, tras sufrir un traumatismo, se someten a una operación de urgencia



Beverley A. Orser

JACK HOLLINGSWORTH *Corbis*

o pacientes en plena cirugía cardíaca; esa dosis inferior puede explicar la aparición de episodios breves de consciencia, como el protagonista del filme.

Aunque sin la aplicación de las actuales técnicas de anestesia general no serían posibles los trasplantes de órganos, ni la cirugía a corazón abierto ni otras operaciones complejas, los poderosos efectos neurodepresivos de esos fármacos hacen que resulten más peligrosos para la vida del paciente que la intervención quirúrgica en sí.

Además, la mortalidad relacionada con la anestesia se ha estabilizado en alrededor de un paciente por cada 13.000 durante los últimos 15 años. Parece, pues, que la anestesiología haya llegado al límite de sus posibilidades. Diversos efectos secundarios graves (desde la pérdida del control de las vías respiratorias hasta dificultades de memoria y cognitivas) se deben también a la acción, tan amplia como desconocida, que los anestésicos modernos ejercen sobre el sistema nervioso central. Ha llegado el momento de dar un paso adelante; las investigaciones más recientes empiezan a alumbrar el camino.

Centro de comunicaciones

Todos los medicamentos que se utilizan para la anestesia general se desarrollaron sólo de forma empírica, es decir, se ensayaron para comprobar su capacidad de producir los efectos que definen el estado anestésico, a saber: sedación, inconsciencia (o hipnosis), inmovilidad, ausencia de dolor (analgesia) y ausencia de recuerdos del período anestésico (amnesia).

Mediante el estudio de los mecanismos por los que operan los agentes anestésicos, numerosos grupos, incluido el nuestro de la Universidad de Toronto, comienzan a desentrañar las bases moleculares y celulares de esos efectos. Los resultados indican que bajo la actividad de esos potentes fármacos subyacen interacciones altamente selectivas con subpoblaciones de células del sistema nervioso; cada interacción específica sería responsable de uno de los efectos de la anestesia.

Esos conocimientos permitirán, al fin, superar la era del éter y desarrollar una nueva generación de fárma-

cos selectivos que se emplearán de forma combinada para producir sólo los efectos deseados y sin riesgo. Además, esas investigaciones están cosechando resultados valiosos para el desarrollo de sedantes, somníferos y otras drogas afines que comparten mecanismos de acción con la anestesia.

Los anestésicos se dividen en dos grandes categorías, según se administren por inhalación, como el isoflurano, o por vía intravenosa, como el propofol. Puede parecer que inducen un sueño profundo, pero el estado que producen la mayoría de los anestésicos generales modernos corresponde más bien a un coma farmacológico. En ese camino hacia la clarificación de los mecanismos subyacentes a los efectos de la anestesia, técnicas como la formación de imágenes por resonancia magnética (RM) y la tomografía por emisión de positrones (TEP) han permitido identificar algunas de las regiones cerebrales y circuitos neuronales que intervienen en cada componente del estado anestésico.

Así, la acción anestésica sobre la médula espinal es responsable de la inmovilidad, mientras que la amnesia se ha vinculado con cambios en el hipocampo, una estructura cerebral que interviene en la formación de la memoria. La disfunción crónica de la memoria que sufren algunos pacientes tras una intervención quirúrgica (uno de los efectos indeseados de la anestesia) puede que sea una secuela de la acción del fármaco sobre el hipocampo.

Dado que la consciencia constituye una experiencia compleja, cuyas propiedades definitorias son aún objeto de debate, no resulta fácil identificar un origen anatómico de la inconsciencia que provoca la anestesia. Una de las teorías más destacadas sostiene que la inconsciencia es mero resultado de una "desvinculación cognitiva", un corte en las comunicaciones entre las regiones del cerebro que cooperan en los procesos cognitivos superiores. A escala local, si imaginamos el entramado de neuronas como las líneas que integran una extensa red telefónica, el efecto de la anestesia general es análogo a la desconexión de las clavijas en la centralita. Pese a la complejidad del fenómeno, los investigadores están logrando progresos alentadores en el descubrimiento del modo en que los anestésicos operan sobre las células del sistema nervioso para bloquear sus transmisiones.

Durante buena parte del siglo XX, se admitió que los anestésicos actuaban mediante la alteración de los componentes líquidos de la membrana celular. La mayoría de los anestésicos corresponden a compuestos altamente liposolubles; su estructura química varía desde la de un simple gas inerte a la de los esteroides más complejos. Tamaña diversidad física y química cimentaba la idea de que los anestésicos rebajaban la actividad neuronal mediante un mecanismo no selectivo.

Sin embargo, la investigación ha terminado por demostrar que esos compuestos interactúan con varios tipos de receptores, unas proteínas específicas que se alojan en la superficie de las neuronas. Cada familia de receptores predomina en áreas concretas del sistema nervioso central. Así pues, la presencia de ciertos tipos de receptor sólo en determinadas subpoblaciones celulares indica sobre qué células opera un anestésico.

Por consiguiente, los estudios actuales se centran en la identificación de los receptores que operan a modo de diana para los anestésicos. También se proponen averiguar la forma en que los fármacos interactúan con los receptores para alterar la función de la célula y el modo en que esos cambios producen los "síntomas", los deseados y los indeseados, de la anestesia.

Receptores GABA

En la superficie de las neuronas se alojan numerosos tipos de proteínas receptoras. Las que han recibido mayor atención en las investigaciones sobre la anestesia son las que se activan mediante neurotransmisores naturales, pues resultan cruciales para regular la transmisión de la información. Los neurotransmisores transmiten mensajes entre neuronas en las sinapsis. Se desplazan desde la neurona presináptica, cruzan un minúsculo intersticio y se unen, por fin, a los receptores de la membrana celular de la neurona postsináptica. Cuando un número suficiente de moléculas neurotransmisoras activa los receptores adecuados, la membrana de la célula postsináptica genera un potencial eléctrico que recorre su longitud hasta la neurona vecina.

Glutamato, serotonina, noradrenalina y acetilcolina son sólo algunos de los neurotransmisores cuya función en la organización del tráfico de señales por el sistema nervioso central se ha estudiando extensamente. No obstante, en la investigación anestesiológica la máxima atención la ha recibido otro neurotransmisor, el ácido gamma-aminobutírico (GABA), por su capacidad de bloquear la comunicación neural.

Neurotransmisor inhibidor, el GABA ayuda a mantener el equilibrio global del sistema nervioso mediante la amortiguación de la capacidad de las neuronas para responder a los mensajes excitadores procedentes de otras moléculas. Se le atribuye

Resumen/*Anestesia selectiva*

- De los anestésicos generales, potentes supresores del sistema nervioso, desconocemos su mecanismo de acción en el cerebro y en el cuerpo.
- El estudio de los mecanismos que subyacen bajo la acción anestésica revela que cada uno de los aspectos del estado anestésico es atribuible a grupos de neuronas, que se caracterizan por las proteínas de superficie (receptores) que interactúan con el fármaco.
- Nuevos compuestos que operasen de forma selectiva sobre receptores determinados y, por tanto, sobre ciertos tipos celulares, se combinarían con sedantes, somníferos y reforzadores de la memoria para ejercer el efecto anestésico deseado.

un papel de primera magnitud en la acción de los anestésicos.

La mayoría de los receptores postsinápticos de las células que interactúan con el GABA corresponden a canales iónicos activados por ligando. Cuando el GABA (el ligando) se enlaza al receptor, éste cambia de conformación: abre de forma temporal un canal que admite la entrada de iones dotados de carga negativa. El aumento en la concentración de iones genera un potencial negativo, que priva a la célula de su capacidad para producir un impulso eléctrico excitador.

El receptor que supuestamente corresponde a la diana principal de los anestésicos es el receptor GABA subtipo A, o $GABA_A$; éste interviene también en la acción de otras drogas sedantes e hipnóticas, sobre todo las benzodiazepinas como el Valium. Concentraciones ínfimas de benzodiazepinas aumentan la función receptora del $GABA_A$. Un simple dato lo confirma: los agentes bloqueadores que obstaculizan la fijación de la benzodiazepina al receptor $GABA_A$ mitigan presto los efectos de esos fármacos.

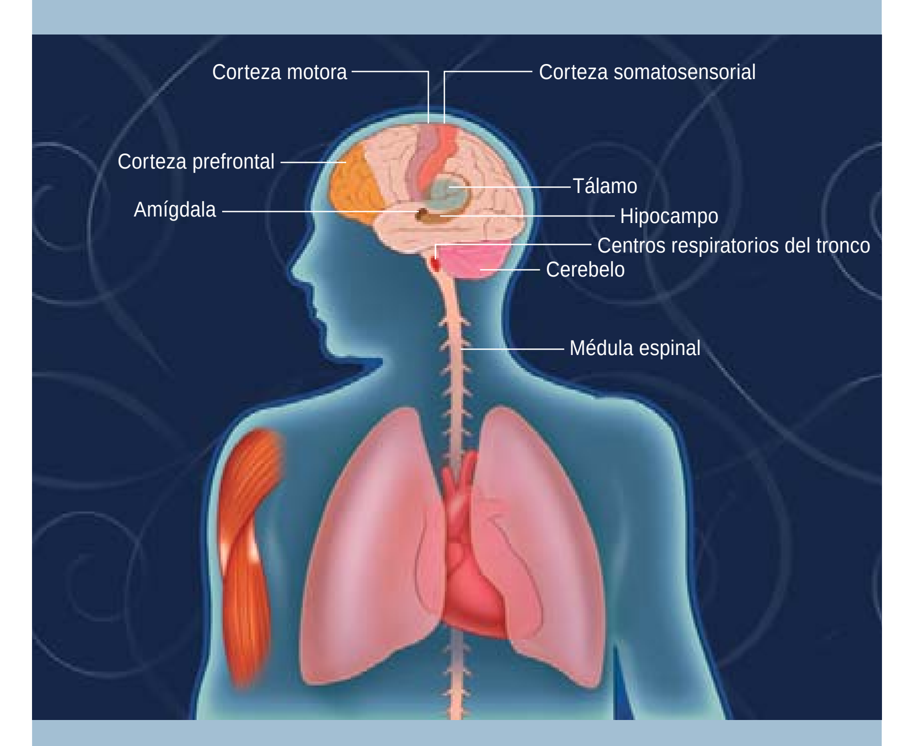
Por desgracia, para los anestésicos generales no existen agentes bloqueadores de ese tipo, que darían pistas respecto a sus receptores diana. No obstante, la investigación sobre cortes de numerosas regiones cerebrales y neuronas obtenidas mediante cultivos histológicos han demostrado que los anestésicos, inhalables e intravenosos, prolongan la duración de las corrientes eléctricas postsinápticas que generan los receptores $GABA_A$.

Se cree que los anestésicos potencian la función de los receptores $GABA_A$ mediante el enlace a cavidades de unión o a aminoácidos concretos de los propios receptores y mediante la prolongación de la abertura del canal, lo cual amplía los efectos inhibidores de las moléculas de GABA fijadas al receptor. A concentraciones elevadas, los anestésicos podrían incluso activar a los receptores GABA solos.

Pero la gran mayoría de las neuronas contienen receptores $GABA_A$. Por eso no se entendía el modo en que los anestésicos influían de forma selectiva, según qué regiones cerebrales. A lo largo de los diez últimos

Efectos generales de la anestesia

Los efectos, deseados e indeseados, de los anestésicos obedecen a su capacidad para suprimir la actividad del sistema nervioso central, que incluye el cerebro y la médula espinal, y controla el ritmo cardíaco y la respiración. Las investigaciones en curso tratan de identificar las estructuras y zonas neurales cuya alteración es responsable de cada uno de los síntomas que definen el estado anestésico.



Componentes del estado anestésico

Sedación
Excitabilidad mermada: tiempos de reacción más largos, habla mascullante y disminución motora. Cae la actividad neuronal en todas las zonas corticales del cerebro.

Inconsciencia (o hipnosis)
Anulación de la percepción y de la respuesta a los estímulos. La depresión cortical es más profunda que en la sedación. Decae también la actividad del tálamo, zona importante para la integración de los procesos cerebrales.

Inmovilidad
Ausencia de motricidad en reacción a estímulos mecánicos o térmicos. La causa de esa parálisis temporal reside en la supresión de la actividad neuronal de la médula espinal, aunque también puede contribuir el cerebelo, zona de control motor.

Amnesia
Ausencia de recuerdos del período anestésico. Los anestésicos inducen cambios en numerosas estructuras cerebrales que intervienen en la formación de la memoria: el hipocampo, la amígdala, la corteza prefrontal y las zonas sensoriales y motoras.

Otros
Relajación muscular y ausencia de dolor (analgesia). Se atribuyen en buena medida a la caída de actividad de la médula espinal.

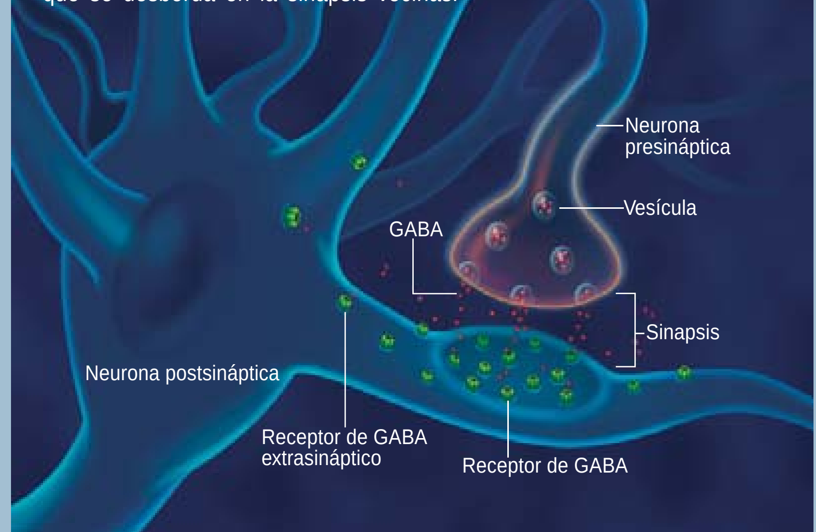
CHRISTINE GRALAPP

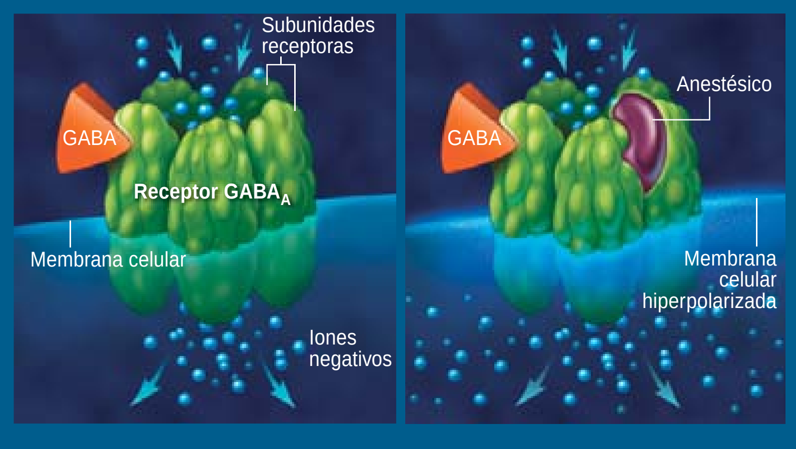
Función de los receptores GABA

Los anestésicos debilitan la comunicación neuronal, en parte, mediante el refuerzo de los efectos del neurotransmisor GABA, una molécula señalizadora que inhibe la activación de las neuronas. La investigación se propone desentrañar la interacción entre fármacos y receptores GABA para bloquear la actividad neural.

Comunicación neuronal

Un impulso eléctrico en la membrana de una neurona provoca la liberación de GABA en la sinapsis, el vínculo con la neurona vecina. Las moléculas atraviesan un pequeño hiato y se unen a receptores GABA alojados en la célula postsináptica. En numerosas zonas cerebrales, existen receptores GABA también fuera de la sinapsis, a lo largo del soma de las neuronas; se activan mediante el GABA que se desborda en la sinapsis vecinas.





Anestésicos y GABA: polarización de la membrana

El subtipo de receptor GABA$_A$ corresponde a un canal que se adentra en la célula postsináptica, compuesto de cinco subunidades proteínicas. Cuando el GABA se une a él, el receptor abre el paso a los iones dotados de carga negativa, con el aumento consiguiente de polarización de la membrana celular y ausencia de generación de un impulso eléctrico (*izquierda*). Se cree que los anestésicos operan mediante la unión a la hendidura del receptor GABA$_A$ y la prolongación de la abertura del canal, lo que provoca la hiperpolarización de la membrana celular (*derecha*).

CHRISTINE GRALAPP

años se han venido registrando, sin embargo, importantes descubrimientos que han puesto de manifiesto que no todos los receptores GABA$_A$ son estructural o farmacológicamente iguales.

El receptor GABA$_A$ consta de cinco subunidades proteínicas, que se mezclan y encajan en varias combinaciones. En los mamíferos hay al menos 19 subunidades de receptor GABA$_A$; la mayoría poseen subtipos, por lo que el número de combinaciones posibles es elevado. Las subunidades que más abundan en las neuronas son, empero, las alfa, beta y gamma. De hecho, la mayoría de los receptores GABA$_A$ se componen de dos subunidades alfa, dos beta y una gamma, aunque a veces, según la región cerebral, en vez de una subunidad gamma hay una delta o una épsilon.

El hallazgo de mayor interés es que las subunidades que componen el receptor definen las propiedades farmacológicas del mismo: una sola variación en una de las subunidades determina la sensibilidad del receptor ante un anestésico determinado.

En cada región cerebral predominan distintos subtipos de receptor GABA$_A$. El estudio de la interacción entre los fármacos y los receptores diana particulares de cada región arroja luz sobre el modo en que los anestésicos producen determinados efectos en distintas partes del sistema nervioso central.

Anestesia selectiva

Nos propusimos desentrañar el mecanismo que subyace bajo la aparición de amnesia. Para ello nos centramos en los receptores GABA$_A$ del hipocampo. Los anestésicos producen amnesia cuando se administra a dosis menores que las necesarias para causar inconsciencia o inmovilidad. Los anestesiólogos lo saben muy bien. Los pacientes rara vez recuerdan las animadas conversaciones que mantienen con ellos mientras los anestesian o salen de ese estado. Con todo, por razones desconocidas, algunos pacientes experimentan un recuerdo inesperado de hechos sucedidos durante la intervención quirúrgica. Por tanto, el hallazgo de los receptores diana correspondientes a los efectos inductores de amnesia quizá facilite la identificación de los pacientes con

riesgo de consciencia intraoperativa, ya que éstos carecen de dichos receptores. Otra posibilidad consistiría en desarrollar estrategias farmacológicas para impedir la consciencia o al menos el recuerdo.

En el curso de nuestro trabajo, descubrimos con sorpresa que incluso los receptores que se alojan fuera de la zona sináptica participaban en la función anestésica. Si la sinapsis opera a modo de centralita, o cuadro de conexiones, en la confluencia de dos células, los receptores alojados en la periferia de la sinapsis o esparcidos a lo largo del soma radican en la misma línea telefónica. Esos receptores $GABA_A$ extrasinápticos son activados incluso por las bajísimas concentraciones de GABA que están presentes de modo natural en el espacio extracelular o que se desbordan en las sinapsis vecinas. Los receptores extrasinápticos abundan en el hipocampo, el tálamo (zona que interviene en la consciencia y en el procesamiento del dolor) y en zonas de la corteza y el cerebelo.

Tras bregar infructuosamente, durante tiempo, en pos de la identificación de los receptores postsinápticos sensibles a las bajísimas concentraciones de anestésico inductoras de amnesia, descubrimos, por una feliz casualidad, la participación de los receptores $GABA_A$ extrasinápticos en la función anestésica.

Habíamos también estado buscando poblaciones de receptores postsinápticos que fueran modulados de forma sinérgica por el midazolam y el propofol, dos de los neurodepresivos intravenosos más usados, sin éxito.

Nuestro trabajo se basaba en el registro electrofisiológico de las corrientes generadas en neuronas de hipocampo en cultivos histológicos. De ese modo, advertimos que las concentraciones de anestésico productoras de amnesia provocaban el aumento de una corriente persistente de baja intensidad generada por receptores $GABA_A$ extrasinápticos. Más que alterar la “centralita”, los fármacos operaban sobre la propia “línea telefónica” mediante la producción de una suerte de zumbido estático o inhibidor que interfería en las comunicaciones.

Descubrimos que el propofol y etomidato, anestésicos inyectables, así como el isoflurano, inhalable, aumentaban hasta 35 veces la intensidad de esa corriente con concentraciones siete veces inferiores a las necesarias para provocar inmovilidad. Otros investigadores, entre ellos Stephen G. Brickley, Mark Farrant y sus colaboradores del Colegio Universitario de Londres, habían ya descrito esa corriente permanente de baja intensidad en ausencia de anestesia. Pero lo que nos sorprendió fue la exquisita sensibilidad que mostraban los receptores extrasinápticos a cantidades ínfimas de anestésicos, inhalables o intravenosos, en comparación con la escasa respuesta que las bajas concentraciones de anestésico causaban en las corrientes postsinápticas. Parecía que los estudios anteriores, entre ellos el nuestro, buscaban la familia de receptores correcta, pero en el lugar equivocado.

Por fin, nuestros experimentos acotaron diferencias estructurales entre los receptores extrasinápticos $GABA_A$ y los receptores sinápticos: éstos carecían de la subunidad alfa-5 que aquéllos solían contener. Parecía que ese pequeño cambio confería una sensibilidad exquisita que permitía detectar cantidades ínfimas de anestésico.

Los resultados concordaban con estudios anteriores. En otros trabajos se había señalado que los receptores $GABA_A$ que portaban la subunidad alfa-5 intervenían en los procesos de memoria normales dependientes del hipocampo, avalando así nuestra teoría de que los receptores extrasinápticos alfa-5 participaban en la aparición de la amnesia.

Para someter a prueba nuestra hipótesis, recurrimos a la experimentación con ratones transgénicos, carentes de la subunidad alfa-5, y ratones silvestres, poseedores del receptor normal. Tal como esperábamos, los múridos silvestres se mostraban sensibles ante las dosis amnésicas de etomidato, mientras de los deficitarios en alfa-5 no manifestaban ninguna alteración de la memoria.

Observamos, asimismo, que la ausencia de receptores $GABA_A$ alfa-5 no repercutía en ninguno de los demás efectos de la anestesia. Sedación, inmovilidad, hipnosis y reacción a los estímulos dolorosos se producían por igual en uno y otro grupo animal.

Esos resultados demostraban que los efectos dañinos del etomidato sobre la memoria eran independientes de otras propiedades de la droga, basadas en la farmacología de subunidades receptoras específicas. Brindaban, además, el primer modelo animal para el estudio de las variaciones en los receptores que pudieran darse entre humanos y explicarían los casos de resistencia a la acción amnésica de algunos anestésicos. Los estudios en curso determinarán si otros fármacos de anestesia general provocan también amnesia mediante la acción preferente sobre los receptores $GABA_A$.

Varios laboratorios europeos y estadounidenses han empleado técnicas experimentales similares para el estudio de los efectos hipnóticos e inmovilizadores de los anestésicos. El grupo dirigido por Gregg E. Homanies, de la facultad de medicina de la Universidad de Pittsburgh, desarrolló un ratón carente de la subunidad delta del receptor $GABA_A$, que confiere una gran sensibilidad a los neurosteroides.

Abordaron la capacidad del fármaco para producir inconsciencia. Y descubrieron que los ratones sin subunidad delta eran menos sensibles a la alfaxalona, un anestésico esteroide. Los ratones mutantes, en cambio, no mostraron diferencia alguna en sus reacciones al propofol, etomidato u otros anestésicos no esteroides, en comparación con el grupo control (silvestre). Aunque no se da un uso generalizado de anestésicos esteroides, los resultados mencionados revelan que sobre subpoblaciones diferentes de receptores $GABA_A$ operan anestésicos distintos.

Los experimentos en cuestión invalidan una idea muy arraigada: habida cuenta de las diferencias químicas entre anestésicos, se admitía, sus múltiples efectos deberían



MAX AGUILERA-HELLWEG

producirse por algún mecanismo no selectivo. El desarrollo empírico de fármacos anestésicos parece haber hallado compuestos que ejercen los mismos efectos, aunque por mecanismos distintos.

Sea por caso el etomidato. Se trata del único anestésico de uso clínico que muestra selectividad para los receptores $GABA_A$ que contienen la subunidad beta-2 o beta-3, pero no la beta-1. La diferencia entre las variantes de subunidad beta sensibles al etomidato y las que no reaccionan responde a un cambio de aminoácido en un punto de la estructura proteínica de la subunidad. La empresa farmacéutica Merck desarrolló ratones transgénicos con una mutación en la ubicación de ese aminoácido de la subunidad beta-2. Descubrió que el efecto hipnótico (pérdida de consciencia) del etomidato era inferior en esos animales; sin embargo, las propiedades inmovilizadoras del fármaco no se alteraban. Uwe Rudolph, entonces en la Universidad de Zúrich, preparó también ratones transgénicos con la misma mutación en la subunidad beta-3. Descubrió que disminuía la acción hipnótica y analgésica del etomidato y el propofol. A modo de contraposición, observó que la alfaxalona mostraba la misma eficacia con los ratones silvestres y los mutantes, prueba de que esas subunidades del receptor no constituían dianas importantes del fármaco.

Seguimos sin saber si las mutaciones puntuales en las subunidades beta-2 y beta-3 influyen en la acción amnésica de los anestésicos. Quedan por descubrir las zonas del sistema nervioso central de los ratones transgénicos que son afectadas por las mutaciones, aunque algunos resultados apuntan hacia los receptores $GABA_A$ extrasinápticos del tálamo. En su conjunto, la investigación atribuye una función destacada al receptor $GABA_A$ en la acción de la anestesia. El paso siguiente consistirá en aplicar a fármacos más selectivos los conocimientos adquiridos a partir de los anestésicos generales.

Tratamientos a medida

Como demuestran el trabajo de mi grupo de investigación y el de

La autora

Beverley A. Orser, profesora de anestesiología y fisiología en a Universidad de Toronto, ejerce en el hospital universitario Sunnybrook.



otros, los receptores extrasinápticos $GABA_A$ alfa-5 del hipocampo resultan esenciales para los efectos inductores de amnesia del etomidato y posiblemente de otros anestésicos generales al uso. De tales resultados se desprende que los fármacos que eluden o tienen como diana ese receptor pasarían por alto, o bloquearían, respectivamente, la formación de memoria.

De hecho, tales compuestos ya se están investigando para otras aplicaciones. En fase de desarrollo preclínico se encuentran medicamentos hipnótico-sedantes que operan sobre la subunidad alfa-5; se supone que carecerán de los efectos amnésicos que provocan los sedantes de benzodiazepina y ciertos somníferos.

Han comenzado los ensayos clínicos del Gaboxadol, el primer fármaco que operará de forma selectiva sobre los receptores $GABA_A$ extrasinápticos para reforzar su función. El Gaboxadol se desarrolló en su origen como antiespasmódico; ahora se investiga su acción inductora del sueño. Dado que actúa sobre los receptores $GABA_A$ con subunidad delta, que se alojan sobre todo en el tálamo y el cerebelo, podría no afectar a la memoria.

El potencial de bloqueo de la memoria de compuestos similares, aunque sin interferir con los receptores alfa-5, resultaría de gran utilidad en las intervenciones quirúrgicas, pues permitiría provocar una amnesia profunda sin abatir la respiración, los reflejos respiratorios o el aparato cardiovascular. Combinado con otros anestésicos, se emplearía un bloqueador de la memoria para prevenir episodios de consciencia intraoperativa. Solo, ese fármaco podría aplicarse al tratamiento de pacientes afectados de trastornos nerviosos postoperatorios mediante la inhibición de recuerdos angustiosos.

El control de los efectos de la anestesia sobre la memoria es sólo una de las aplicaciones de una nueva anestesiología basada en fármacos dirigidos. En numerosas situaciones, resulta innecesaria e indeseable la profunda neurodepresión que causan los anestésicos al uso. Mediante un cóctel de fármacos, cada uno productor de un efecto, la anestesia del futuro permitiría que un paciente departiera mientras le reparan una extremidad o permaneciera inmóvil y sedado, aunque despierto, mientras le sustituyen una cadera. Ese enfoque multifarmacológico ya se aplica en el tratamiento del dolor postoperatorio.

Nunca ha sido la anestesia más segura que hoy, aunque no se halle exenta de riesgos. Tenemos ante nosotros la oportunidad de superar la era del éter y encaminarnos hacia un modelo moderno de la asistencia anestésica.

Bibliografía complementaria

ANESTHESIA SAFETY: MODEL OR MYTH? Robert S. Lagasse en *Anesthesiology*, vol. 97, págs. 1609-1617; diciembre, 2002.

MOLECULAR AND NEURONAL SUBSTRATES FOR GENERAL ANESTHETICS. Uwe Rudolph y Bernd Antkowiak en *Nature Reviews Neuroscience*, vol. 5, págs. 709-720; septiembre, 2004.

EMERGING MOLECULAR MECHANISMS OF GENERAL ANAESTHETIC ACTION. Hugh C. Hemmings y otros en *Trends in Pharmacological Sciences*, vol. 6, n.º 10, págs. 503-510; octubre, 2005.

$\alpha 5GABA_A$ RECEPTORS MEDIATE THE AMNESTIC BUT NOT SEDATIVE-HYPNOTIC EFFECTS OF THE GENERAL ANESTHETIC ETOMIDATE. Victor T. Cheng y otros en *Journal of Neuroscience*, vol 26, n.º 14, págs. 3713-3720; 5 de abril, 2006.

Autoorganización de la vegetación

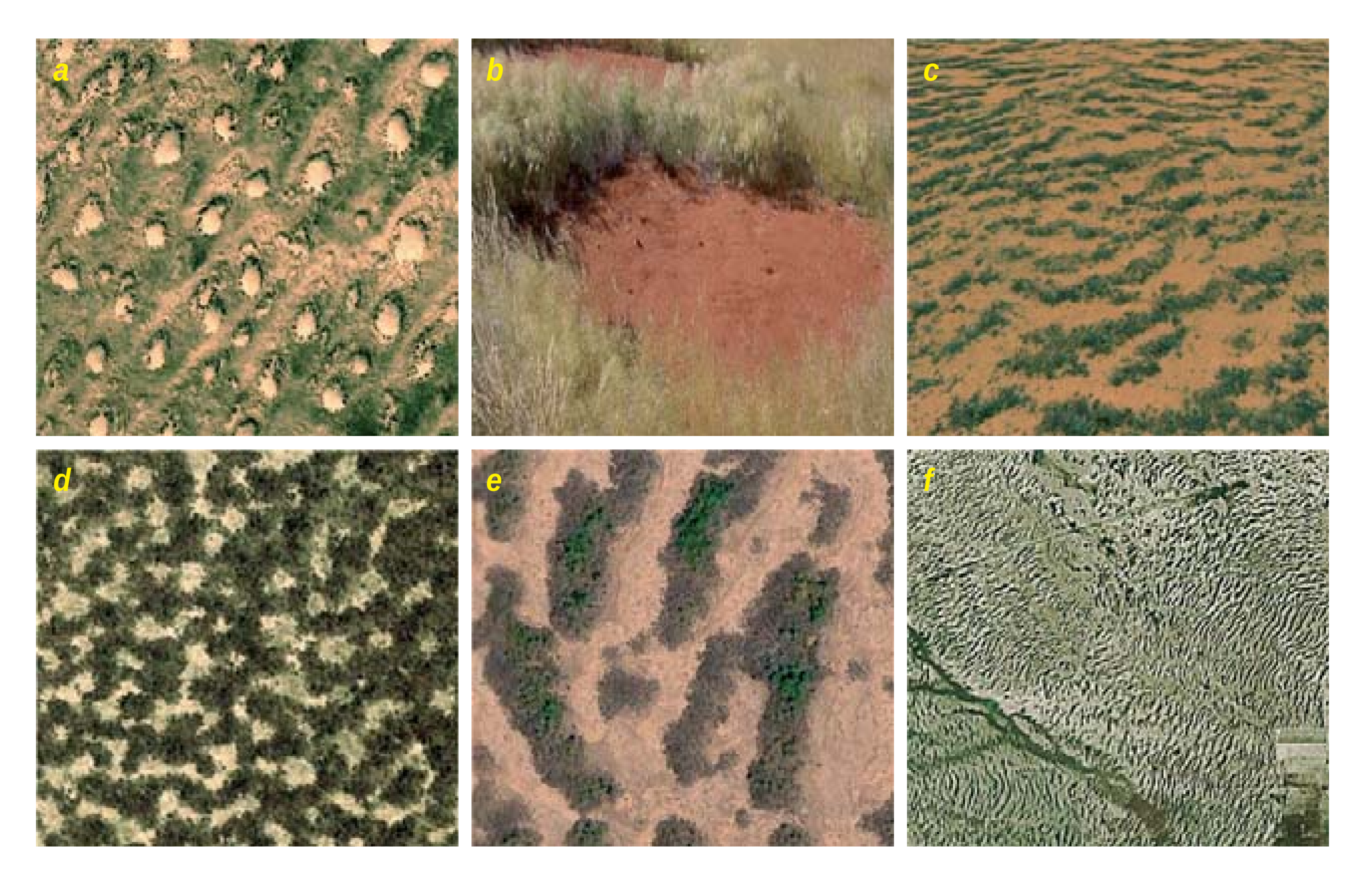
en tierras áridas y semiáridas

Las zonas áridas y semiáridas conforman aproximadamente el 30 por ciento de las superficies emergidas. No son raros en ellas los paisajes donde la vegetación crea mosaicos. A vista de pájaro aparecen manchas vegetales alternando con otras sin vegetación. Cada paisaje aparece dominado por una determinada especie vegetal y un tamaño característico de las manchas, que pueden ir desde un pequeño matojo (0,5-2 m²) hasta grandes masas arbóreas (100-1000 m²) de mulga (*Acacia aneura*) o bandas de varias hectáreas de hierba de Mitchell (*Astrebla pectinata*), por citar dos ejemplos australianos.

Ni la distribución espacial ni la forma de las manchas son aleatorias; en muchas ocasiones se disponen sobre el terreno creando patrones geométricos muy dispares: bandas paralelas, círculos regularmente espaciados, estructuras en panal y otras muchas conformaciones. Encontramos idénticas formaciones en lugares tan distantes entre sí como Norteamérica, el Sahel o Australia (*véase la figura 1*).

Las precipitaciones en las regiones de que hablamos son estacionales, en forma de fuertes chaparrones, muy irregulares, y con alternancia de largas sequías y años húmedos. El mosaico es la respuesta a este peculiar régimen de lluvias: así lo ponen de manifiesto las observaciones de campo, que señalan que los mosaicos funcionan como sistemas hidroecológicos, integrados y fuertemente acoplados, con flujos verticales y horizontales e interacciones en todas las escalas de tiempo y espacio.

Detrás de esta compleja caracterización se esconde un funcionamiento básico sencillo. Durante un chaparrón los espacios sin vegetación, con escasa infiltración, se saturan pronto de agua, mientras que no es así en los vegetados, con mayor capacidad de infiltración a consecuencia de la actividad biológica en el suelo. Como resultado se genera un flujo superficial que transporta recursos (agua, nutrientes y sedimentos) de los primeros a los segundos; de ese modo, éstos alcanzan un nivel de producción



1. Algunas formas regulares en mosaicos vegetales: toma aérea (*a*) y detalle (*b*) de lunares en un fondo de vegetación herbácea en Marienfluss, Namibia; (*c*) paisaje rayado de *Acacia aneura* en la vertiente norte de los montes Macdonnell, cerca de Alice Springs, Australia; (*d*) estructura laberíntica y ejemplo de *brouse tigrée* (*e*), o matorral atigrado, en el suroeste de la República del Níger; (*f*) paisaje rayado en el desierto de Chihuahua, México.

2. Fotografía de la *brousse tigrée* saheliana, donde se indica la productividad biológica en las distintas zonas. La zona intersetos, desnuda de vegetación, está recubierta de una fina corteza cementada que no facilita la infiltración, lo que provoca durante una precipitación una escorrentía en la dirección de la pendiente (*flecha blanca*) que es interceptada por el seto de la izquierda. El agua se infiltra en éste de forma diferencial, dependiendo directamente de la densidad de biomasa: cuanto mayor es la biomasa, tanto mayor es la capacidad de infiltración. Como la biomasa es mayor en el centro del seto, allí se capturará más agua y se registrará una productividad mayor. Esta última puede multiplicar de 100 a 400 veces la obtenida en la zona intersetos.

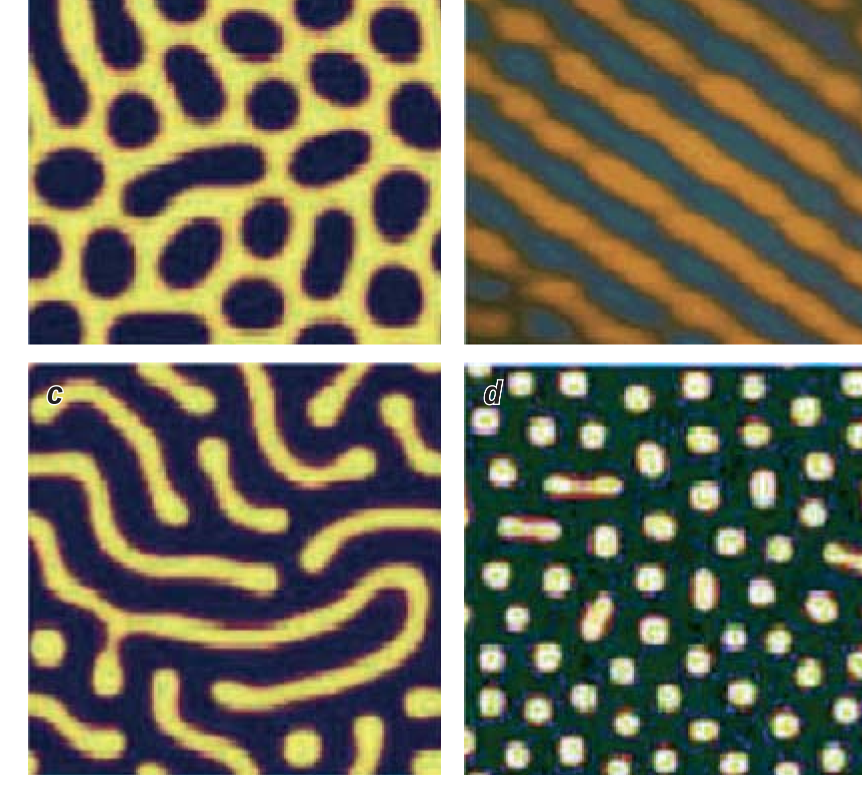


biológica muy por encima de lo que permitirían las condiciones climáticas (*véase la figura 2*).

Ante la escasez de recursos, la naturaleza "adopta", pues, la solución de concentrarlos en ciertas zonas del terreno en detrimento de otras. ¿Por qué? La regularidad encontrada en formaciones como las de la figura 1 induce a pensar que tal estructuración del paisaje pudiera depender de una causa exógena. El orden vendría "impuesto desde fuera". De hecho, una matriz de dunas fósiles, por ejemplo, sería el factor externo que imprime las bandas de vegetación.

Pero la explicación más satisfactoria apela a razones puramente endógenas. Es el propio paisaje el que se autoorganiza. Para entenderlo debemos recurrir a las inestabilidades de Turing, así llamadas en honor del matemático inglés que en los años cincuenta del pasado siglo atribuyó la diferenciación morfogenética a la difusión. Este fenómeno, esencialmente homogeneizador, puede tener el efecto contrario si se dan ciertas condiciones de retroalimentación de ciertos agentes morfogenéticos del sistema. Turing habló en abstracto y no especificó la naturaleza de estos "morfógenos", que hemos de individuar en cada sistema al que se aplique su teoría.

Si se dan las condiciones de Turing, el estado homogéneo da paso a una "inhomogeneización" espacial, no aleatoria, hacia formas geométricas simples, que surgen como resultado exclusivo de la interacción entre elementos del propio sistema. El mecanismo de Turing se concreta matemáticamente en un sistema de ecuaciones diferenciales en derivadas



3. Ejemplos de estructuras producidas en una reacción química a consecuencia de inestabilidades de Turing. Se ha escogido una coloración arbitraria para marcar contrastes. Aunque más regulares que las de las formas exhibidas en la figura 1, saltan a la vista las similitudes, que ponen de manifiesto la universalidad de la teoría de Turing. Las irregularidades en las estructuras de Turing observadas en paisajes áridos o semiáridos son debidas a la complejidad adicional introducida por la tipología y topografía del terreno, que no tiene paralelo en una reacción química en laboratorio.

parciales y éstas, a su vez, sirven de modelo en otros muchos contextos. Es decir, las condiciones de Turing pueden manifestarse en otros ámbitos: la química, la física o la biología tienen abundantes ejemplos en los que acontece la inestabilidad prevista por Turing (*véase la figura 3*).

Vistas las sorprendentes semejanzas entre las formas mostradas a raíz de una inestabilidad de Turing y los mosaicos vegetales de la figura 1, resulta inmediato pensar en un paralelismo formal, con mecanismos apropiados de retroalimentación para los mosaicos. ¿Cuáles serían esos mecanismos? La respuesta es muy simple: estriba, por una parte, en la relación directa entre la capacidad de absorción de agua del suelo y la frondosidad de la cubierta vegetal, y, por otra, en los pulsos de crecimiento de biomasa en respuesta a la captación de aguas de escorrentía —producidas por los parches desnudos— en zonas de vegetación.

Esta teoría explica también la formación de los distintos patrones. Escojamos, por ejemplo, un terreno sin pendiente y hagamos simulaciones con niveles crecientes de precipitación. Para valores muy bajos (menos de 150 mm anuales), la ausencia total de vegetación es un estado estable, hasta que se alcanza un valor umbral, que podríamos cifrar en 200 mm, en el que el estado se desestabiliza y se origina un conjunto de manchas, más o menos circulares, regularmente espaciadas. Para una precipitación mayor, esta última tipología se desestabiliza a su vez, dando paso a una disposición "nudosa", como la de las figuras 1d y 3c. La secuencia de inestabilidades continúa hasta llegar a un paisaje vegetal homogéneo. Otro caso es el de una ligerísima pendiente (para evitar la erosión producida por la escorrentía). Aparecerá un paisaje en bandas vegetales a lo largo de las curvas de nivel de la pendiente. En el Sahel, recuerdan el moteado del tigre, de ahí su designación local de *brousse tigrée* (*véase figura 1e*).

La naturaleza muestra muchas regularidades geométricas universales. Los paisajes naturales no escapan a la regla. Por lo tanto, no es de extrañar que un mismo marco teórico, las inestabilidades de Turing, resulte satisfactorio en ámbitos tan distintos como la química y las formaciones vegetales en zonas áridas y semiáridas. Las segundas ocurren, claro está, en áreas remotas, lejos de las alteraciones humanas, lo que explica lo poco conocidas que son. Sin embargo, en ellas habremos de encontrar los modelos que permitan apropiadas restauraciones ambientales, con un aprovechamiento óptimo de recursos, como alternativa a las fracasadas repoblaciones indiscriminadas.

VÍCTOR FAIRÉN LE LAY
Departamento de Física Matemática y Fluidos,
UNED, Madrid

Comunicación celular

Proteínas adaptadoras endocíticas

La morfología, el crecimiento y la homeostasis de los tejidos en organismos pluricelulares dependen de una estricta regulación de la respuesta a los factores de crecimiento. Mediante la unión de estos factores (o ligandos) a sus respectivos receptores se forma en la célula un amplio repertorio de complejos proteicos altamente dinámicos. De ese modo, la señal se amplifica y se traduce en cambios en los estados del proteoma y genoma celulares.

Distintos ligandos pueden unirse, a la vez, a una misma célula para formar complejos multiproteicos. Un solo ligando provoca múltiples respuestas, en función de cómo se formen estos complejos de señalización o *señalosomas*. Se desconocen por el momento la especificidad de la acción de los señalosomas, la regulación de su localización y vida media, el modo en que se propaga la señal desde la membrana plasmática al núcleo y otros aspectos esenciales de la regulación de estas respuestas a los factores de crecimiento. Sin embargo, se sabe que la endocitosis (o internalización) reviste suma importancia en dichos procesos.

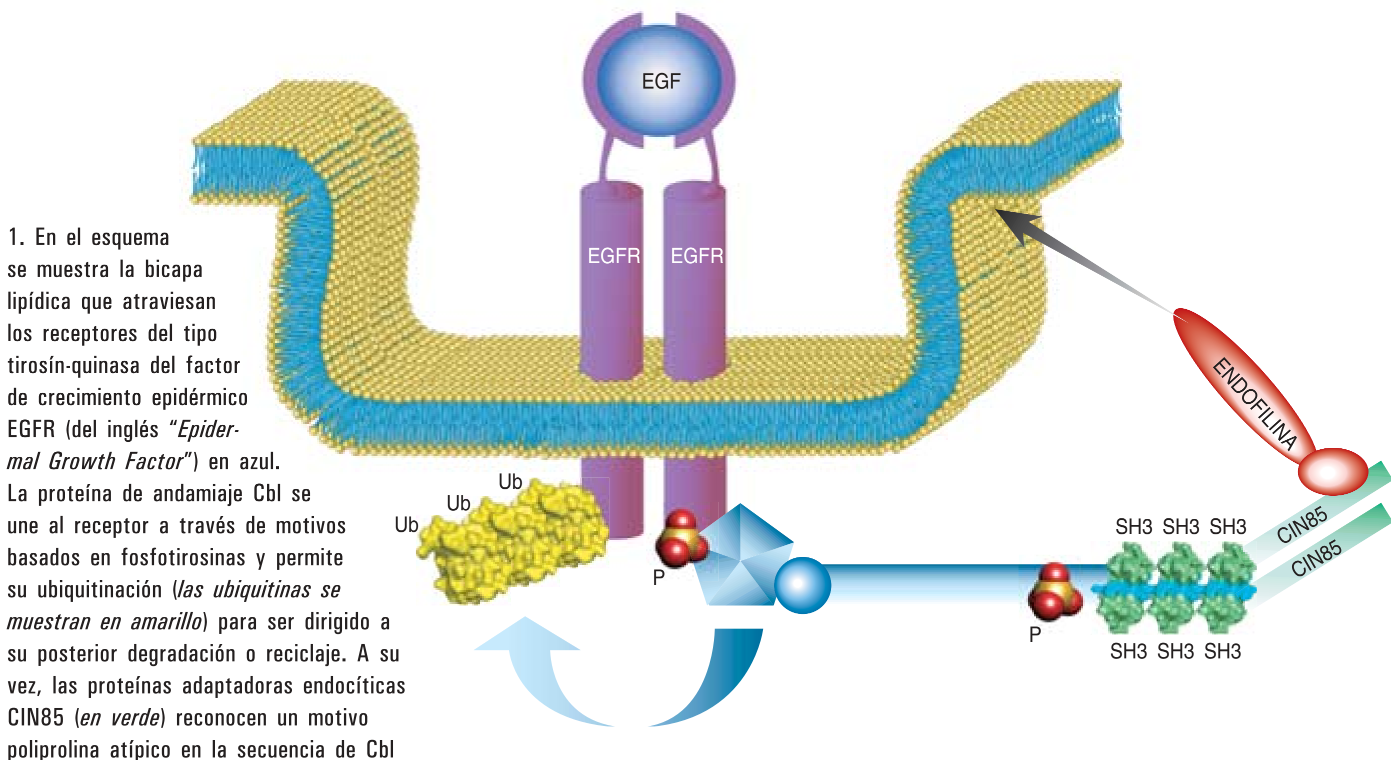
Las proteínas de membrana no pueden redistribuirse en el citosol. Por ello, la regulación de su concentración en la membrana celular requiere la internalización y degradación de las mismas. En el caso de los receptores celulares, la endocitosis resulta imprescindible para lograr una insensibilización prolongada de las células a los factores de crecimiento y otros estímulos externos. Los complejos multiproteicos controlan la maquinaria endocítica mediante modificaciones postraduccionales como la fosforilación y la ubiquitinación; de ese modo, influyen en el destino del *señalosoma*.

Se ha demostrado que la interrupción de la endocitosis causa alteraciones en la duración y la especificidad de los acontecimientos desarrollados en este proceso. Por ello, se ha propuesto que las alteraciones en el mecanismo natural de la endocitosis podrían ser de carácter oncogénico.

Por otro lado, la membrana endosómica constituye un mosaico dinámico de dominios lipídicos que acoplan distintos conjuntos de proteínas endocíticas, conectadas entre sí por proteínas adaptadoras. Las moléculas adaptadoras corresponden a polipéptidos, sin función catalítica en su mayoría, dotados de uno o varios dominios que se unen a otros ligandos (proteicos o no proteicos).

Las moléculas adaptadoras controlan de forma selectiva el ensamblaje espacial y temporal de los complejos multiproteicos, transmisores de las señales intracelulares que regulan el crecimiento, la diferenciación, la migración y la supervivencia celulares. Existe un subgrupo de moléculas adaptadoras, las proteínas de andamiaje (Cbl y CIN85, por ejemplo), que se acoplan a complejos macromoleculares mediante la unión simultánea entre sí y con otras proteínas.

La alteración en las señales que clasifican los receptores de membrana para su degradación contribuye a la iniciación y progresión de algunas neoplasias. Esto sucede a menudo en el contexto de la activación constitutiva de receptores de factores de crecimiento como los receptores de tirosín-quinasa (RTK), implicados en cánceres humanos y en modelos experimentales. Por ejemplo, la desviación de la regulación negativa mediada por Cbl constituye un fenómeno recurrente en la desregulación oncogénica de los RTK. Una vez que



1. En el esquema se muestra la bicapa lipídica que atraviesan los receptores del tipo tirosín-quinasa del factor de crecimiento epidérmico EGFR (del inglés "*Epidermal Growth Factor*") en azul. La proteína de andamiaje Cbl se une al receptor a través de motivos basados en fosfotirosinas y permite su ubiquitinación (*las ubiquitinas se muestran en amarillo*) para ser dirigido a su posterior degradación o reciclaje. A su vez, las proteínas adaptadoras endocíticas CIN85 (*en verde*) reconocen un motivo poliprolina atípico en la secuencia de Cbl a través de sus dominios SH3, y facilitan la endocitosis del receptor mediante la unión a otras proteínas como la endofilina (*en naranja*).

el ligando se ha unido al receptor, con su estimulación consiguiente, la proteína Cbl se une al receptor a través de motivos basados en fosfotirosinas. El receptor sufre luego una ubiquitinación, que le dirige a su posterior degradación o reciclaje. A su vez, las proteínas adaptadoras endocíticas CIN85 reconocen un motivo poliprolina atípico en la secuencia de Cbl a través de sus dominios SH3, facilitando así la endocitosis del receptor mediante la unión a otras proteínas como la endofilina.

Conviene señalar que los dominios SH3 constituyen un ejemplo clásico de módulos de interacción proteica. Se conoce ya en detalle el tipo de interacciones que establecen. Sin embargo, el reconocimiento de Cbl por los dominios SH3 de CIN85 no corresponde a ningún modo de reconocimiento descrito con anterioridad. Recientemente, se ha descubierto la formación de un complejo ternario en el que dos dominios SH3 de CIN85 se unen de forma simultánea al mismo motivo derivado de Cbl en orientaciones opuestas. Se ha demostrado, además, que la unión simultánea de los dominios resulta necesaria para una correcta internalización y degradación de receptores tirosín-quinasa.

A pesar de los avances mencionados, para comprender los mecanismos de regulación implicados en el desarrollo celular es necesario ahondar todavía más en las interacciones moleculares que tienen lugar en la internalización de los receptores de membrana. El conocimiento de dichas interacciones podría sentar las bases para el futuro descubrimiento de nuevas terapias antineoplásicas.

2. Dos dominios SH3 de CIN85 (*verde*) unidos al mismo péptido derivado de la secuencia de Cbl-b (*azul*) en orientaciones opuestas. (Las coordenadas atómicas se han obtenido del fichero 2BZ8 de la base de datos del *Protein Data Bank*, publicado en *Nature Structural & Molecular Biology*, vol. 12, n.º 11, págs. 972-979; 2005.)

NAYRA CÁRDENES ESTELLÉS
JERÓNIMO BRAVO SICILIA
Centro Nacional
de Investigaciones Oncológicas,
Madrid

Alícia Duró y Francisco Amat

Biodiversidad en las salinas

1. Ecosistema salino donde medra *Artemia*.

2. Cuando las condiciones ambientales son desfavorables, los embriones adoptan un estado de vida quiescente (diapausa); se rodean de una envoltura (corion) que les protege de las agresiones externas. Se forman así quistes de entre 200 y 300 micras de diámetro.

El buque oceanográfico *Glomar Challenger* (1970) encontró, a 150-200 metros bajo la superficie del fondo marino, unas capas de sal de una milla de potencia. Son restos de las salmueras que llenaron el mar Mediterráneo cuando se secó, hace 5,5 millones de años, y quedó reducido a una serie de lagos salados en sus zonas más profundas. Se supone que fue entonces cuando apareció *Artemia*, un crustáceo branquiópodo anostráceo que medra en aguas hipersalinas.

En 1756 D. Schlosser describió taxonómicamente el género *Artemia* a partir de unos ejemplares procedentes de las salinas de Lymington. En la clasificación sistemática que Carl Linné realizó en 1758, el crustáceo se denominó *Cancer salinus*. En 1819, W. E. Leach, suponiendo que se trataba de un género monoespecífico, lo reclasificó como *Artemia salina*. En la actualidad se estima que existen 6 especies bisexuales de artemias y un variado grupo de estirpes partenogenéticas de ploidía diversa.

Dada su condición de artrópodo, *Artemia* crece según un proceso de metamorfosis. De forma periódica, muda el exoesqueleto y aumenta el tamaño. Al tiempo que se desarrollan las partes del cuerpo, se diferencian las antenas y los ojos pedunculados en la cabeza, los filópodos en el tórax y el ovisaco del aparato reproductor en el abdomen de la hembra. En condiciones favorables ese proceso dura dos semanas. Al final, aparece un individuo adulto, macho o hembra, de entre 7 y 8 milímetros de talla.

Artemia corresponde a un organismo filtrador que se nutre de las partículas en suspensión en el medio acuático. Se vale para ello de sus 11 pares de filópodos torácicos, dotados de un constante movimiento metacrónico. Las partículas retenidas suelen ser microalgas y arquibacterias halófilas (que proliferan en medios hipersalinos). El mismo mecanismo le facilita la natación y la respiración.

Se trata de un crustáceo de distribución cosmopolita, excluida la Antártida. Sin embargo, el actual uso irrestricto de *A. franciscana* en acuariofilia y acuicultura ha facilitado su introducción en Europa y Asia, donde se ha convertido en especie exótica invasora, desplazando y eliminando a las autóctonas. Ello, junto con el cambio climático, constituye una enorme amenaza a la biodiversidad de este organismo.

3. Cuando las condiciones ambientales vuelven a ser adecuadas, los quistes se rehidratan y los embriones reemprenden el desarrollo. En menos de 24 horas se forman los nauplios, que miden entre 400 y 600 micras, según la especie.







4. El macho adulto (*arriba*) es menor que la hembra (*derecha*), que cuenta con un ovisaco donde se acumulan los ovocitos. El macho presenta 2 penes retráctiles; se distingue por unos ojos pedunculados y unas antenas hiperdesarrolladas, que le sirven para aferrarse a la hembra durante la fecundación.

FOTOGRAFIAS DE JUAN CARLOS NAVARRO. MICROSCOPIA CON TINCION DE LUGOL (*arriba*); IMAGENES OBTENIDAS CON LUPA BINOCULAR (*centro y abajo*)



5. En cada ciclo reproductivo hembra y macho se acoplan. Este introduce uno de los dos penes por la abertura del ovisaco de aquélla. El esperma llega hasta los óvulos maduros. La fecundación da origen a la formación de nauplios (reproducción ovovivípara) o de quistes de resistencia (reproducción ovípara).

De las partículas a la cosmología

La historia de la cosmología de partículas, nueva rama de la física que intenta dilucidar los orígenes del universo, demuestra que la ciencia puede beneficiarse con los cambios de rumbo

David Kaiser

La cosmología de partículas, que investiga el efecto de las unidades más pequeñas de la materia en la forma y el destino del universo, es una de las disciplinas más activas de la física actual. En estos últimos años, el Departamento de Energía, la Fundación Nacional para la Ciencia de Estados Unidos y la NASA la han subvencionado con unos quinientos millones de dólares. Se han dado grandes pasos en el conocimiento de las interacciones de las partículas de altas energías que rigieron el universo en los primeros momentos de su historia. Lo que entonces sucedió influiría en la evolución cósmica de los miles de millones de años subsiguientes.

El éxito impresionante de la cosmología de partículas sorprende aún más cuando se tiene en cuenta que ni siquiera existía 30 años atrás. Antes de 1975 se consideraba (especialmente en los EE.UU.) que la física de partículas y la cosmología eran campos de estudio separados; pocos se detenían a pensar en cómo una especialidad podría ayudar a la otra.

¿Por qué, pues, surgió la cosmología de partículas? A mediados de los años setenta, algunos comprendieron que el universo primitivo ofrecía un observatorio sin par sobre los fenómenos de energía demasiado elevada para que se los reconstruyese en el laboratorio. Pero una serie de cambios en la financiación y la enseñanza de la física contribuyeron también a situar las cuestiones cosmológicas en primer plano. La rapidez con que emergió la cosmología de partículas ilustra hasta qué punto los presupuestos gubernamentales, las instituciones educativas e incluso los libros de

LOS COSMOLOGOS DE PARTICULAS investigan los primeros momentos de la historia cósmica.



SLIM FILMS

texto pueden alterar radicalmente la dirección de las investigaciones científicas. La historia de esa época demuestra, asimismo, que la ciencia puede cosechar enormes beneficios cuando los investigadores se alejan de los temas trillados para abordar nuevos retos.

Una buena manera de relatar esa historia consiste en centrarse en la fortuna de dos conjuntos de ideas: el campo de Brans-Dicke, concebido por especialistas en la gravitación, y el campo de Higgs, planteado por físicos de partículas. Ambos grupos crearon esos conceptos en respuesta a un problema candente entre los físicos de finales de los años cincuenta y principios de los sesenta: ¿por qué tienen masa los objetos? Aunque esas dos hipótesis no promovieron la unión de la física de partículas con la cosmología, el curso que siguieron nos enseñará el proceso de convergencia de ambas ramas de la física.

Una historia de dos φ

La masa parece una característica de la materia tan obvia, que podría pensarse que no requiere explicación. Pero describir la masa de manera que concuerde con otras ideas de la física moderna no resultó fácil. Los expertos en gravitación y cosmología enmarcaban el problema dentro del principio de Mach. Ernst Mach, físico y filósofo austriaco, fue crítico famoso de Newton e inspirador del joven Albert Einstein. Una buena enunciación aproximada del principio de Mach podría ser ésta: la masa de un objeto —un parámetro que mide su resistencia a cambiar de movimiento— deriva en última instancia de las interacciones gravitatorias que tenga con toda otra materia del universo. Aunque este principio interesó a Einstein y estimuló su pensamiento, la teoría general de la relatividad se aparta del mismo.

Para incorporar el principio de Mach en la teoría gravitatoria se presupuso la existencia de un nuevo campo escalar que interaccionaba con todos los tipos de materia. (Un campo escalar viene dado en cada punto del espacio y del tiempo por un solo número.) En 1961, Carl Brans, doctorando de la Universidad de Princeton, y su director de tesis, Robert Dicke, señalaban que, en la relatividad general de Einstein, la intensidad de la gravedad viene fijada por la constante de Newton, G.

Según Einstein, G tiene el mismo valor en la Tierra que en las galaxias más remotas y no cambia con el tiempo. A modo de exposición alternativa, Brans y Dicke sugirieron que el principio de Mach se satisfaría si la constante de Newton variase en el tiempo y en el espacio. Introdujeron un campo llamado φ inversamente proporcional a la constante de Newton e intercambiaron G por $1/\varphi$ en las ecuaciones gravitatorias de Einstein.

Según la teoría de Brans-Dicke, la materia responde a la curvatura del espacio y el tiempo, como en la relatividad general ordinaria, *y* a las variaciones de la intensidad local de la gravedad. El campo φ impregna todo el espacio; su comportamiento contribuye a determinar el movimiento de la materia a través del espacio y del tiempo. Cualquier medición de la masa de un objeto depende, por lo tanto, del valor local de φ. Esta teoría era tan atractiva, que los miembros del grupo de investigaciones gravitatorias de Kip Thorne, del Instituto de Tecnología de California, bromeaban diciendo que creían en la relatividad general de Einstein los lunes, miércoles y viernes y en la gravedad de Brans-Dicke los martes, jueves y sábados. (Los domingos eran agnósticos ante una y otra.)

Mientras tanto, en la comunidad, mucho mayor, de los físicos de partículas el problema de la masa presentaba una forma diferente. Ya en los años cincuenta, se vio que podrían representarse los efectos de las fuerzas nucleares imponiendo determinadas clases de simetrías en las ecuaciones que gobernaban el comportamiento de las partículas subatómicas. Sin embargo, los términos que habrían de incluirse en dichas ecuaciones para representar las masas de las partículas violaban las simetrías exigidas.

En particular, la aporía afectaba a los bosones W y Z, las partículas que dan lugar a la fuerza nuclear débil, responsable de la desintegración radiactiva. Si estas partículas transmisoras de las fuerzas careciera n verdaderamente de masa, como las simetrías parecían requerir, el alcance de las fuerzas nucleares debería ser infinito: dos protones, por ejemplo, ejercerían su fuerza nuclear mutua a través de la galaxia. Un alcance tan grande contradecía flagrantemente el comportamiento observado de las fuerzas nucleares, que disminuyen rápidamente en cuanto las distancias son mayores que el tamaño de los núcleos atómicos. Sólo si las partículas que transmiten las fuerzas tenían masa podía concordar el alcance teóricamente predicho con las observaciones.

Muchos físicos abordaron el rompecabezas, afanosos por dar con una teoría que, a la vez, presentara las simetrías de las fuerzas subatómicas e incluyese partículas con masa. En 1961 Jeffrey Goldstone, entonces en la Universidad de Cambridge, observó que las soluciones de las ecuaciones no tenían por qué obedecer las mismas simetrías que las ecuaciones. Como simple ilustración, introdujo un campo escalar, por coincidencia designado también φ, cuya densidad de energía potencial, $V(\varphi)$, tocaba fondo en dos puntos: cuando φ tenía los valores $-v$ y $+v$. Como la energía del sistema toma el valor más bajo en estos mínimos, el campo acabará en uno de ellos. La energía potencial es exactamente igual para ambos

Resumen/*Una revolución en la física*

- Hasta los años setenta, se consideraba que la física de partículas y la cosmología eran campos de estudio independientes.
- La merma de medios que la física de partículas sufrió desde finales de los años sesenta incitó a los científicos a ampliar sus horizontes y acercarse a la gravitación y la cosmología.
- En los años ochenta se vio que el estudio del universo primitivo aportaba nuevas perspectivas para la exploración de las altas energías. Desde entonces, el campo híbrido de la cosmología de partículas se ha convertido en uno de los más fructíferos de la física.

valores; ahora bien, puesto que el campo debe acabar en uno de los dos valores —sea *–v* o *+v*—, la solución de las ecuaciones romperá espontáneamente su simetría.

En 1964, Peter W. Higgs, de la Universidad de Edimburgo, reconsideró el trabajo de Goldstone y halló que una teoría con rotura espontánea de simetría permitía la existencia de partículas portadoras de masa. La masa surge de las interacciones entre el campo φ y todos los tipos de partículas, incluidas las que generan la fuerza nuclear débil. Las ecuaciones que gobiernan estas interacciones, demostró Higgs, obedecen los requisitos de simetría. Antes de que φ se sitúe en uno de los mínimos de su energía potencial, las partículas se mueven sin demasiados obstáculos. Sin embargo, una vez que φ llega a *+v* o *–v*, ofrece resistencia a lo que se le acople: el equivalente subatómico a moverse en una melaza. De ese modo, las partículas transmisoras de la fuerza (y también la materia corriente, los electrones, digamos) se comportan como si tuvieran una masa distinta de cero; cualquier medida de su masa dependerá del valor local de φ.

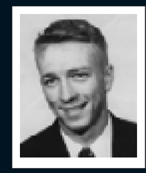
Los trabajos de Brans-Dicke y los de Higgs se publicaron casi al mismo tiempo y en la misma revista, *Physical Review*. Ambos se difundieron rápidamente. Y se cuentan entre los artículos de física más citados de todos los tiempos. Los dos explicaban el origen de la masa por medio de un nuevo campo escalar que interaccionaba con todos los tipos de materia.

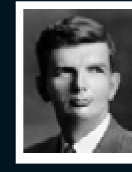
Dada la semejanza de las propuestas y la atención que de inmediato se les prestó, cabría pensar que quien tratase de una trataría también de la otra. Pero ese emparejamiento rara vez se hizo. De los 1083 artículos que citaron los trabajos de Brans-Dicke o de Higgs entre 1961 y 1981, sólo seis —menos de un 0,6 por ciento— incluyeron ambos artículos en sus referencias. (El primero de esos seis artículos apareció en 1972; los otros cinco, después de 1975.) Ese mutuo desconocimiento pone de manifiesto las rígidas separaciones que existían en aquel entonces entre los físicos de partículas y los especialistas en gravitación y cosmología.

Distintos conceptos de masa

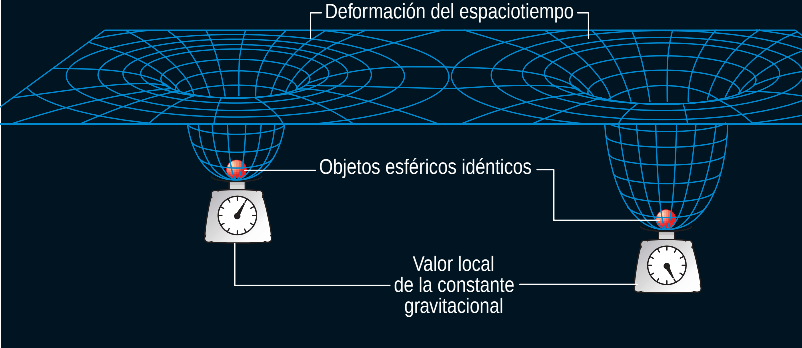
Las barreras que dividían a los físicos a principios de los años sesenta se reflejan en sus intentos paralelos de explicar por qué los objetos tienen masa. Aunque los cosmólogos y los físicos de partículas propusieron teorías similares, pocos vieron la conexión entre ellas.

Desde la cosmología: gravedad de Brans-Dicke


Carl Brans


Robert Dicke

▼ En 1961, Carl Brans y Robert Dicke, de la Universidad de Princeton, concibieron un campo llamado φ que permitía que la constante gravitacional de Newton variara en el tiempo y en el espacio. Un objeto en un punto del espacio en donde la constante sea pequeña (*izquierda*) tendrá menos masa —y curvará menos el espaciotiempo local— que un objeto idéntico en un punto donde la constante sea grande (*derecha*).

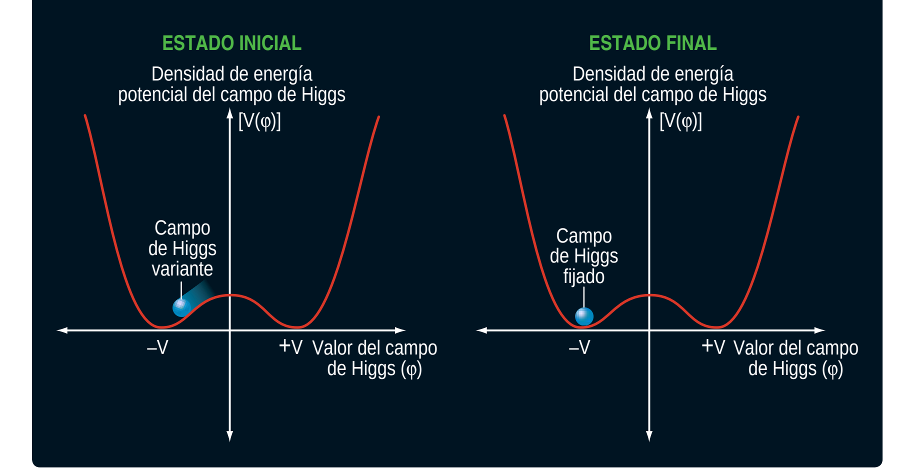


Desde la física de partículas: el campo de Higgs

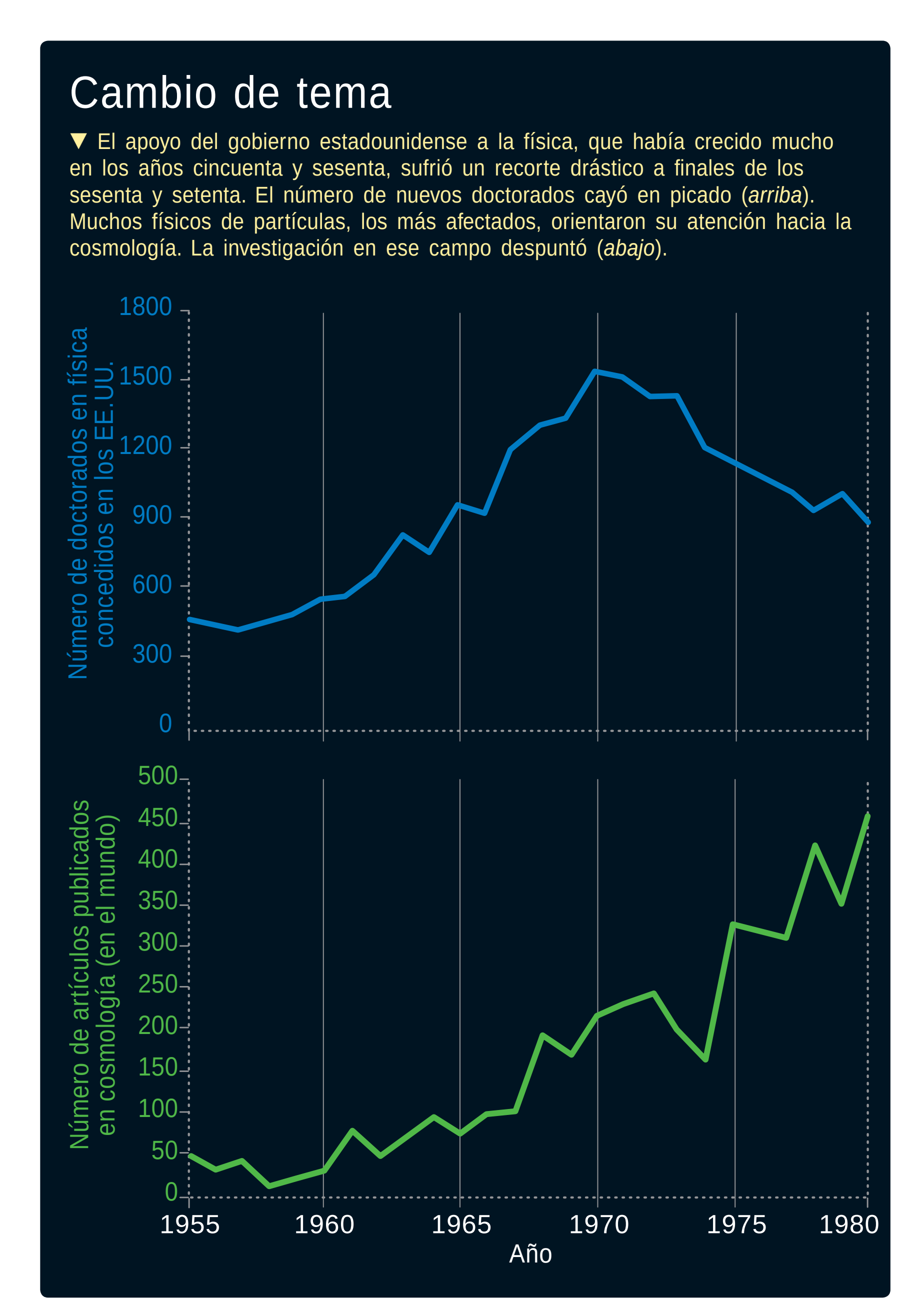

Jeffrey Goldstone


Peter Higgs

▼ En 1961, Jeffrey Goldstone, entonces en la Universidad de Cambridge, introdujo un campo también llamado, por casualidad, φ, cuya densidad de energía potencial V(φ) tiene un mínimo en dos puntos: *-v* y *+v*. Tres años después, Peter Higgs, de la Universidad de Edimburgo, utilizó este campo para explicar la masa. Las partículas no tenían masa al principio, cuando φ variaba (*izquierda*); adquirieron masa sólo después de que φ alcanzara uno de sus mínimos (*derecha*).



LUCY READING-IKKANDA (*diagramas de campo*); R. CRESSON *Luno* (*Brans*); BETTMANN/CORBIS (*Dicke*); AIP EMILIO SEGRE VISUAL ARCHIVES, PHYSICS TODAY COLLECTION (*Goldstone*); PETER TUFFY *Universidad de Edimburgo* (*Higgs*)



Cambio de tema

▼ El apoyo del gobierno estadounidense a la física, que había crecido mucho en los años cincuenta y sesenta, sufrió un recorte drástico a finales de los sesenta y setenta. El número de nuevos doctorados cayó en picado (*arriba*). Muchos físicos de partículas, los más afectados, orientaron su atención hacia la cosmología. La investigación en ese campo despuntó (*abajo*).

LUCY READING-IKKANDA

Crisis y transición

Las dos comunidades vieron diferentes cosas en sus respectivos φ. A los expertos en gravitación y cosmología, el campo de Brans-Dicke φ_{BD} les entusiasmó porque ofrecía una posibilidad alternativa a la relatividad general de Einstein. A los físicos de partículas, el campo de Higgs φ_H les atrajo porque les daba esperanzas de que sus teorías llegarían a explicar las fuerzas nucleares entre partículas dotadas de masa. Hasta mediados de los años setenta, nadie imaginó que esos φ_{BD} y φ_H fueran físicamente similares, ni siquiera que mereciese la pena examinarlos al alimón.

La división entre la física de partículas y la cosmología era especialmente tajante en EE.UU. cuando Brans, Dicke, Goldstone y Higgs formularon sus respectivos φ. Por botón de muestra: el Comité de Revisión de la Física, de la Academia Nacional de Ciencias, publicó en 1966 un informe que recomendaba doblar los fondos y las plantillas de los doctorados de la física de partículas durante los años siguientes; en cambio, proponía una expansión casi nula de las áreas, ya reducidas de por sí, de la gravitación, la cosmología y la astrofísica.

Algunos libros de texto soviéticos de física gravitatoria publicados por entonces especulaban sobre las fuerzas nucleares; tal mezcla de especialidades no se daba nunca en las obras norteamericanas análogas.

Sin embargo, tales pautas cambiarían radicalmente a finales de los años setenta. Siempre se señala que un par de avances estimuló la fusión de los dos campos: el descubrimiento de la libertad asintótica en 1973 y la construcción de las primeras teorías de gran unificación, o GUT, en 1973 y 1974. La libertad asintótica remite a un inesperado fenómeno de ciertas clases de teorías que gobiernan las interacciones de partículas: la intensidad de la interacción disminuye a medida que la energía de las partículas aumenta. (En cambio, la intensidad crece en la mayoría de las demás fuerzas.)

Por primera vez, los físicos de partículas se sintieron capacitados para acometer cálculos precisos y fiables sobre la interacción nuclear fuerte —que mantiene a los quarks ligados dentro de las partículas nucleares, como protones y neutrones, partículas nucleares—, a condición de que se restringieran a energías muy altas, mucho más altas que las que se habían investigado experimentalmente.

La introducción de las GUT dirigió, asimismo, la atención hacia las muy altas energías. Se vio que las intensidades de las tres fuerzas fundamentales —el electromagnetismo y la interacción nuclear débil y la interacción nuclear fuerte— iban convergiendo a medida que crecían las energías de las partículas. Se planteó la hipótesis de que, a partir de cierta energía muy alta, las tres fuerzas actuarían como una fuerza única, indiferenciada. La escala de energías en la que aparecería esa gran unificación era inmensa: del orden de 10^{24} electronvolt, o más de un billón de veces mayor que las máximas energías de los aceleradores de partículas.

La escala de energías de las GUT no se podría alcanzar nunca en laboratorios terrestres, pero si el universo entero había comenzado con una “gran explosión” caliente, la energía media de sus partículas tuvo que ser altísima en los primeros momentos de la historia del cosmos.

Con el advenimiento de la libertad asintótica y las GUT, a los físicos de partículas se les presentó una razón

obvia para estudiar el universo primitivo: los primeros momentos de la gran explosión proporcionaban el "acelerador de los pobres" que permitiría observar las interacciones de alta energía que no se podían recrear en la Tierra. Científicos, periodistas, filósofos e historiadores han sostenido que ahí se encuentra la explicación de que apareciera la cosmología de partículas.

Pero, ¿eso es todo? Pese a la importancia innegable de los avances de la teoría de partículas, no bastaban para explicar el crecimiento del nuevo subcampo. Por un lado, la sincronía no es perfecta. Las publicaciones sobre cosmología (en todo el mundo, así como en los EE.UU.) venían ya siendo mucho más abundantes desde antes de 1973; no afectó lo más mínimo al ritmo al que crecía la publicación de trabajos sobre la libertad asintótica y las GUT. Por otra parte, las GUT no merecerían mucha atención, ni siquiera entre los teóricos de partículas, hasta finales de los años setenta y principios de los ochenta. Tres de las primeras revisiones generales del estado de la cuestión del campo emergente de la cosmología de partículas, publicadas entre 1978 y 1980, ignoraron la libertad asintótica y las GUT.

Las nuevas ideas, por sí solas, no bastaron para preparar el camino de la cosmología de partículas; cambios gubernamentales y educativos desempeñaron también un papel importante. Hasta mediados los años sesenta, los físicos de EE.UU. se habían beneficiado de la "burbuja de la guerra fría", un período en el que el gobierno federal prodigó fondos en educación, defensa e investigación científica. Pero a finales de esa década, las reducciones drásticas ocasionadas por las protestas contra la guerra del Vietnam, el deshielo de la guerra fría y la Enmienda Mansfield, que restringía mucho el gasto del Departamento de Defensa en investigación básica, causaron estragos en la física estadounidense.

Casi todos los campos de la ciencia y de la ingeniería entraron en recesión, pero la física bajó más deprisa y más profundamente que cualquier otra disciplina. El número de nuevos doctores en física se desplomó, cayendo casi tan rápidamente de 1970 a 1975 como había aumentado durante los años posteriores al Sputnik.

A partir del decenio de 1960, bruscos recortes causaron estragos en la física en los EE.UU.

La financiación federal de la física mermó. Perdió más de un tercio (en dólares constantes) entre 1967 y 1976. Desde los años cincuenta hasta mediados de los sesenta, el número de puestos de trabajo disponibles siempre había sido mayor que el número de estudiantes de física que buscaban trabajo en el servicio de empleo del Instituto Americano de Física. Pero las perspectivas laborales empeoraron rápidamente: 989 aspirantes compitieron por 253 puestos en 1968, 1053 por 53 en 1971.

La física de partículas fue, con mucho, la más perjudicada; entre 1970 y 1974, los fondos federales que recibía se quedaron en la mitad. Comenzó un rápido éxodo de talentos: de 1968 a 1970, el número de investigadores que abandonó en EE.UU. la física de partículas duplicó la cifra de quienes se incorporaron a ella. El número de nuevos doctores en física de partículas cayó un 44 por ciento entre 1969 y 1975, el descenso más rápido de cualquier rama de la física.

Mas, al propio tiempo, la suerte empezó a sonreírles a la astrofísica y la gravitación. En parte por el estímulo de una serie de hallazgos importantes de mediados de los años sesenta —los cuásares, los púlsares, el fondo cósmico de la radiación de microondas—, el número de nuevos doctores en esa área creció un 60 por ciento entre 1968 y 1970 y en otro 33 por ciento entre 1971 y 1976, a pesar de que el número total de doctores en física bajó en picado.

Para analizar el desastre, el Comité de Revisión de la Física publicó en 1972 un nuevo informe que resaltaba las dificultades que padecía la física de partículas. A muchos teóricos jóvenes formados en esa disciplina, observaba el comité, les estaba costando cambiar a otras materias debido a su "acotada especialización".

El informe urgía a los departamentos de física a modificar la formación de los teóricos de partículas: "Los grupos de las universidades tienen una responsabilidad en poner en contacto a sus más brillantes y capaces alumnos con las oportunidades que brindan todas las especialidades de la física".

No tardaron en introducirse cambios en los planes de estudio de las universidades, encaminados hacia un mejor conocimiento de otras áreas de la física por los estudiantes graduados; se hizo mayor hincapié en la gravitación y la cosmología. Por todo el país, los programas de física empezaron a ofrecer nuevos cursos. Tras haber ignorado la gravitación y la cosmología durante decenios, los editores estadounidenses publicaron numerosos libros de texto sobre esos temas para satisfacer una demanda repentina.

Inflación

Estos precipitados cambios dejaron su marca en la interpretación de los campos de Brans-Dicke y de Higgs. En 1979, luego de casi veinte años sin que apenas hubiere quien mencionase los dos campos en un mismo trabajo, no digamos ya quien los considerase físicamente análogos, dos teóricos norteamericanos, y de forma independiente, sugirieron que φ_{BD} y φ_H pudieran ser uno y el mismo campo. Anthony Zee, entonces en la Universidad de Pennsylvania, y Lee Smolin, entonces en la de Harvard, unieron las dos piezas clave de φ combinando las ecuaciones gravitatorias de Brans-Dicke con un potencial de Goldstone-Higgs con rotura de simetría. (Teóricos de fuera de los EE.UU. habían expuesto ideas parecidas entre 1974 y 1978, pero no se les prestó mucha atención en aquellos años.)

En ese modelo, la intensidad local de la gravedad variaba inicialmente en el espacio y el tiempo, con G proporcional a $1/\varphi^2$, pero su actual valor constante se alcanzaba cuando el campo φ llegaba a un mínimo del potencial con rotura de simetría, lo cual se suponía que había sucedido en los primeros momentos de la gran explosión. De esta manera, Zee y

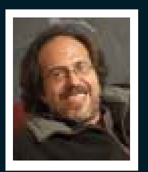
La conexión

A finales de los años setenta, una nueva generación de físicos, familiarizados a la vez con la teoría de partículas y con la cosmología, percibió la relación entre la gravedad de Brans-Dicke y el campo de Higgs.



ANTHONY ZEE

Durante la carrera, Zee trabajó con el experto en gravitación John Wheeler en la Universidad de Princeton. Después se doctoró en teoría de partículas. Renovó su interés por la cosmología en París, durante su año sabático de 1974.



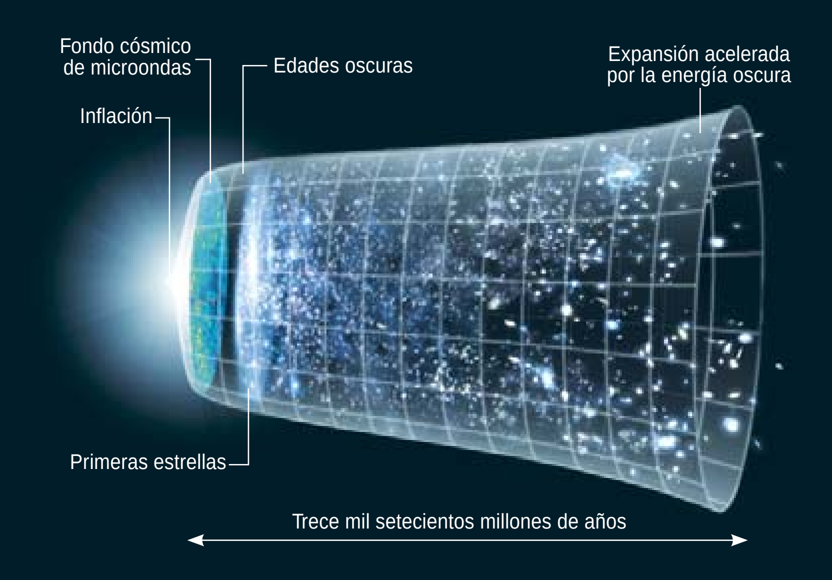
LEE SMOLIN

En los años setenta, Smolin estudió teoría de partículas y cosmología, siendo estudiante graduado, en la Universidad de Harvard. También trabajó con Stanley Deser, de la Universidad Brandeis, pionero de la gravedad cuántica.



ALAN GUTH

Guth se doctoró en física de partículas por el Instituto de Tecnología de Massachusetts en 1972. Se interesó por la cosmología tras asistir a una conferencia de Dicke a finales de los años setenta.



▲ En sendos artículos publicados en 1979, Zee y Smolin combinaron las ecuaciones gravitatorias de Brans-Dicke con un potencial con rotura de simetría de Goldstone-Higgs. En 1981 Guth introdujo otro campo, basado en el de Higgs, el inflatón. Este campo proporcionaba la fuerza impulsora de un hipotético período de expansión supercélere —o inflación— acaecido durante los primeros instantes del universo.

Smolin ofrecieron una explicación de la debilidad de la fuerza gravitatoria, comparada con las demás fuerzas: cuando el campo llega a su estado final, $\varphi = \pm v$, φ queda fijado en un cierto valor grande, distinto de cero, y, por ende, G (inversamente proporcional a v^2) adquiere un valor pequeño.

Las trayectorias profesionales de Zee y de Smolin ilustran el acercamiento de los físicos a la cosmología, una vez reventó la burbuja de la guerra fría. Zee había trabajado de estudiante con el experto en gravitación John Wheeler en Princeton, a mediados de los años sesenta; luego se doctoraría en teoría de partículas en Harvard. Obtuvo el título en 1970, justo cuando la disciplina en que iba a especializarse empezaba a sufrir mermas. Como más tarde recordaría, ni siquiera se hizo mención de la cosmología en su período de graduación. Terminada su etapa posdoctoral, empezó a enseñar en Princeton. Pasó un año sabático, 1974, en París. En el piso que le alquiló a un físico francés encontró muchos trabajos de teóricos europeos que intentaban explicar propiedades cosmológicas mediante la teoría de partículas (por ejemplo, que haya en el universo observable más materia que antimateria). Aunque no le parecieron convincentes, reavivaron su pasado interés por la gravitación. Al volver del año sabático, y de nuevo en contacto con Wheeler, reorientó su actividad hacia la cosmología de partículas.

Lee Smolin, en cambio, empezó su doctorado en Harvard en 1975, cuando los cambios curriculares entraron en vigor. Aparte de la teoría de partículas en que iba a graduarse, estudió gravitación y cosmología y fue estrecho colaborador de Stanley Deser (de la cercana Universidad Brandeis), profesor visitante en el departamento de física de Harvard. Deser era uno de los pocos teóricos norteamericanos que habían procurado formular en los años sesenta una descripción de la gravitación compatible con la mecánica cuántica. Fue el primero que publicó, en 1972, un artículo que citase a la vez a Brans-Dicke y a Higgs (aunque trataba los dos campos de manera un poco diferente y en partes distintas). Smolin, que investigaba la gravedad cuántica, propuso, en 1979, a punto de terminar su tesis, que φ_{BD} y φ_H pudieran ser el mismo campo.

La experiencia de Smolin es característica de la generación de teóricos a la que pertenece, la de los formados entre mediados y finales de los años setenta. Paul Steinhardt, Michael Turner y Edward "Rocky" Kolb estudiaron gravitación y teoría de partículas en el doctorado. Pronto, Smolin, Turner, Kolb, Steinhardt y otros formarían a sus propios alumnos de doctorado para que trabajasen en la nueva área híbrida de la cosmología de partículas. Para estos jóvenes teóricos y su creciente nú-

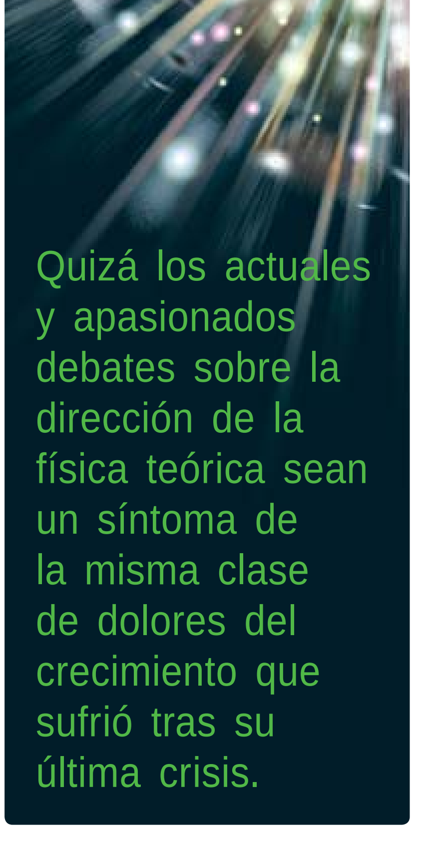
mero de alumnos, era natural asociar φ_{BD} y φ_H. Tanto Turner como Kolb y Steinhardt dirigieron grupos de investigación que buscaron, a lo largo de los años ochenta, nuevas relaciones entre φ_{BD} y φ_H.

Basándose en su trabajo de 1979, Zee observó en 1980 que las teorías cosmológicas ordinarias —así, la gran explosión— no explicaban la extraordinaria regularidad del universo observable (al menos a las mayores escalas). Por su parte, Dicke llegó a la conclusión de que la gran explosión tampoco explicaba la geometría plana del universo, pese a que la forma de éste podía, en principio, diferir de la curvatura mínima observada por los astrónomos.

En 1981 Alan H. Guth, becario posdoctoral de la Universidad de Stanford, hoy profesor del Instituto de Tecnología de Massachusetts, recurrió a la cosmología inflacionaria para abordar ambos problemas. El modelo de Guth se cimentaba en otro campo escalar, parecido al de Higgs. Apodado inflatón, el campo en cuestión aportaba la fuerza impulsora de una supuesta expansión supercélere, o "inflación cósmica", acaecida en los primeros instantes del universo.

La trayectoria profesional de Guth recuerda la de Zee. Terminó su doctorado en teoría de partículas en el Instituto de Tecnología de Massachusetts en 1972, antes de las extensas reformas curriculares que llevaron de nuevo la gravitación a las aulas estadounidenses. El derrumbe de la física de partículas afectó de lleno a Guth. En los años siguientes ocupó diversos puestos posdoctorales. A finales de los años setenta asistió a una conferencia de Dicke sobre el problema de la planitud del universo. De allí salió con la idea de que la cosmología podría resultar de interés en lo que se refería a los problemas sin resolver de la teoría de partículas.

Sumergido en la nueva física de las GUT, y mientras se esforzaba por dominar los fundamentos de la gravitación y la cosmología, dio con la inflación. La mayor parte de los físicos que desarrollaron la idea, sin embargo, eran jóvenes teóricos —entre ellos Steinhardt, Kolb, Turner y sus alumnos— a los que ya se había preparado para un empeño semejante.



Andrei Linde, entonces en el Instituto Lebedev de Física de Moscú, estaba igualmente predispuesto a explorar las ideas inflacionarias: habiendo estudiado en Rusia, donde la física de partículas y la gravitación hacía tiempo que iban de la mano, no tardaría en ampliar la teoría.

Desde entonces, se ha convertido en una rutina para los cosmólogos de partículas el combinar los campos de Brans-Dicke, de Higgs y del inflatón, y adaptar libremente las ecuaciones para explicar fenómenos muy dispares. Este salto conceptual pasó de impensable a imperceptible por acostumbrado en sólo unas cuantas generaciones académicas. Tamaño cambio de actitud demuestra el poder de la pedagogía y la inmensa influencia que los cambios institucionales ejercen en el pensamiento científico.

¿Puede la historia repetirse? La física de partículas sufrió de nuevo en los años noventa (sobre todo con la anulación del Colisionador Superconductor, un poderoso acelerador de partículas que se estaba construyendo en Texas). Desde entonces, los fondos han seguido disminuyendo en EE.UU. Quizá las apasionadas discusiones de hoy en día sobre la dirección de la física teórica, entre los defensores de la teoría de cuerdas y quienes abogan por otros enfoques, sean un síntoma de la misma clase de dolores del crecimiento que padeció tras la anterior crisis.

Se esperan los resultados que manarán de los proyectos que se inaugurarán el año que viene: el Gran Colisionador de Hadrones en Suiza, el Telescopio Espacial de Rayos Gamma de Zona Amplia y el satélite Planck, que medirá el fondo cósmico de microondas con una exactitud sin precedentes. Con suerte, la física de altas energías saldrá de nuevo a flote, tan vibrante, tan fascinante como hace 30 años.

El autor

David Kaiser es físico e historiador. Recibió su doctorado en física teórica y en historia de la ciencia en la Universidad de Harvard y ahora es profesor asociado del programa de ciencia, técnica y sociedad del Instituto de Tecnología de Massachusetts y docente en el departamento de física de ese mismo centro.

Bibliografía complementaria

Was Einstein Right? Putting General Relativity to the Test. Segunda edición. Clifford M. Will. Basic Books, 1993.

The Inflationary Universe: The Quest for a New Theory of Cosmic Origins. Alan Guth. Addison-Wesley, 1997.

Einstein's Universe: Gravity at Work and Play. Anthony Zee. Oxford University Press, 2001.

Three Roads to Quantum Gravity. Lee Smolin. Basic Books, 2001.

Cold War Requisitions, Scientific Manpower, and the Production of American Physicists after World War II. David Kaiser en *Historical Studies in the Physical and Biological Sciences*, vol. 33, págs. 131-159; 2002.

Inflationary Cosmology: Exploring the Universe from the Smallest to the Largest Scales. Alan Guth y David Kaiser en *Science*, vol. 307, págs. 884-890; 11 de febrero, 2005.

Kaushik Basu

El dilema



MATT COLLINS

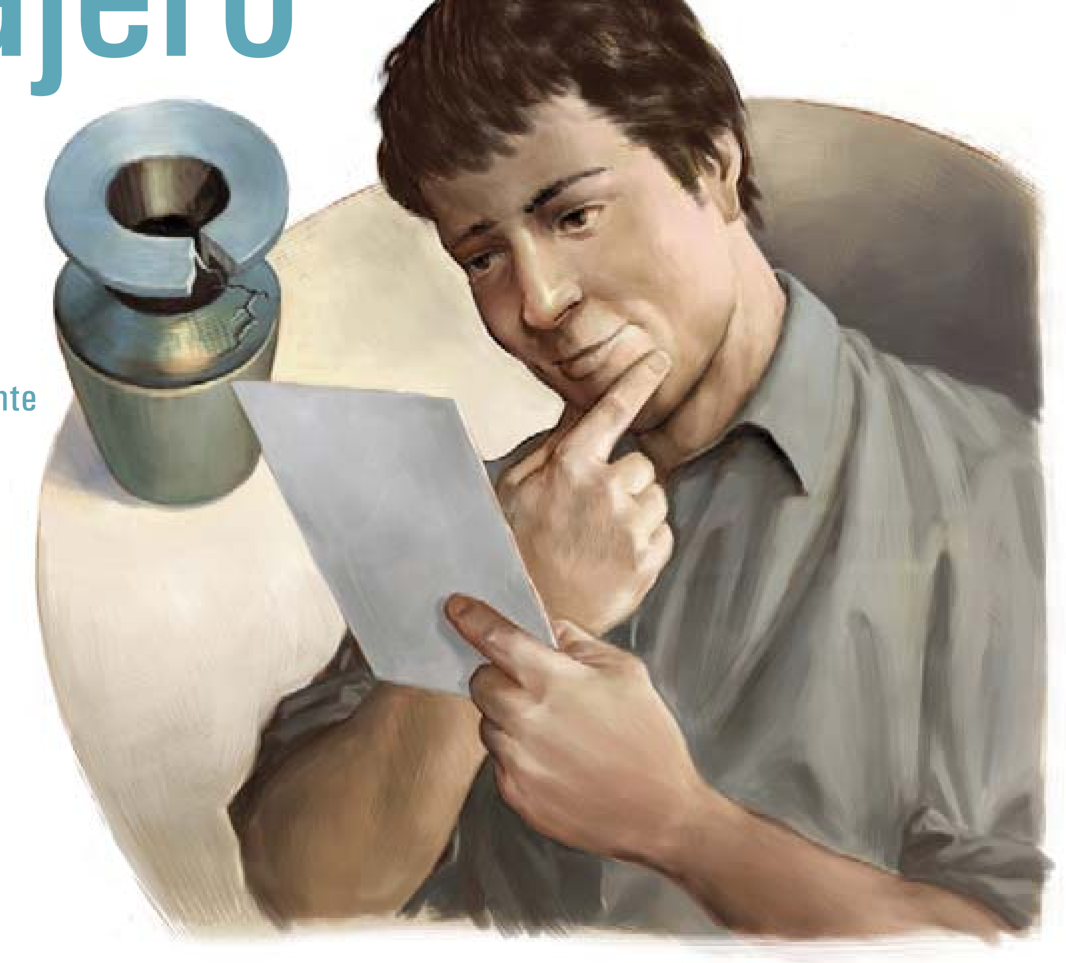
Luci y Pablo, que acaban de regresar de una isla perdida del Pacífico, se dan cuenta de que en el avión se han deteriorado los recuerdos, idénticos los de ella y los de él, que adquirieron allí. Un representante de la compañía se muestra dispuesto a indemnizarles, pero ignora el valor de esos raros objetos. No le parece sensato preguntárselo sin más a los reclamantes, porque seguramente van a exagerar.

Idea, pues, un plan más complicado. Le pide a cada uno de ellos que ponga por escrito el valor que les asigna, un número entero de euros comprendido entre 2 y 100. Los reclamantes ya no pueden comunicarse entre sí. Si ambos declaran el mismo número, el funcionario aceptará que ése ha sido el precio auténtico, y se lo abonará sin discusión. Pero —les informa— si los números de una y otro son diferentes, supondrá que el verdadero valor es el menor de los dos e indemnizará a los dos reclamantes con la menor de las dos cantidades, aunque con una bonificación y una penalización: la reclamación más baja será premiada con 2 euros adicionales, por su honradez; mientras que la petición más elevada sufrirá una penalización de 2 euros. Por ejemplo, si Luci había reclamado 46 y Pablo había exigido 100 euros, Luci recibirá 48 euros y Pablo, 44 euros.

¿Qué cantidades van a solicitar Luci y Pablo? ¿Cuánto reclamaría usted?

Los supuestos de esta clase, en los que uno o varios individuos han de tomar decisiones que serán remuneradas en función de la opción elegida, reciben el nombre de juegos por quienes las estudian (los profesionales de la “teoría de juegos”). El autor ideó éste, al que llamó “Dilema del viajero”, en 1994, con varios objetivos en mente: poner en tela de juicio las miopes concepciones de la conducta racional y los procesos cognitivos que han abrazado muchos economistas y politólogos; criticar los supuestos de la teoría económica tradicional propios del liberalismo clásico, y poner de relieve una paradoja lógica de la racionalidad.

del viajero



En este sencillo juego, los participantes rechazan sistemáticamente la opción tenida por racional, pero al actuar de forma ilógica acaban cosechando una remuneración mayor. Tal resultado exige un razonamiento formal de nuevo tipo

El dilema del viajero (DV) logra estos fines porque la lógica del juego dicta que la mejor opción consiste en reclamar 2 euros, mientras que casi todo el mundo elige un valor próximo a 100 euros; y así lo hacen tanto quienes no han efectuado consideraciones lógicas de ninguna clase como quienes comprenden perfectamente que se están desviando por completo de la opción tenida por "racional". Los jugadores consiguen una recompensa mucho mayor no adhiriéndose a la razón en este caso. Hay, pues, algo racional cuando se opta por no serlo en el juego del DV.

El juego, en los años transcurridos desde que fue concebido, ha adquirido vida propia, con investigadores que lo han generalizado y sometido a experimentos de laboratorio. Tales estudios han arrojado luz sobre los procesos de decisión humanos. Quedan, no obstante, cuestiones pendientes sobre la forma en que pueden aplicarse la lógica y el razonamiento al DV.

Nash y el sentido común

Para comprender por qué la elección lógica es 2, examinemos las consideraciones que podría hacerse Luci. Su primera idea fue la de escribir el mayor valor posible, 100, que le proporcionaría 100 euros si Pablo fuese tan codicioso como ella. (Si el objeto en cuestión hubiese costado mucho menos de 100 euros en realidad, Luci estaría felicitándose por lo ingenuo del plan del funcionario.) Pero a Luci no tarda en ocurrírsele que si en lugar de 100 escribiera 99, podría ganar un poco más, pues si Pablo reclamase 100 euros, ella ganaría 101 euros. Pero eso mismo ha debido ocurrírsele a Pablo, y si ambos solicitasen 99 euros, eso recibirían ambos. Así que si Pablo escribiese 99 euros, a ella le convendría más escribir 98, pues entonces obtendría 100 euros. Pero el mismo razonamiento llevaría a Pablo a solicitar 98 euros; de ser así, ella debería rebajar su petición a 97, para poder ganar 99. Y así sucesivamente.

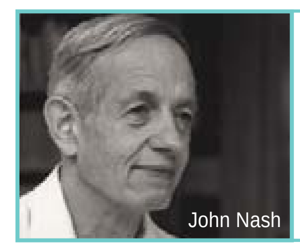
De seguir esta cadena de razonamientos, ambos acabarían solicitando el mínimo permisible, a saber, 2 euros. Puede parecer sumamente inverosímil que Luci llegue a rebajar su petición hasta 2 euros a causa de un razonamiento así, aunque no es eso lo que importa (de hecho, es justamente su inverosimilitud lo que aquí se sostiene). Pero es adonde nos conduce cierta clase de lógica.

El tipo de análisis expuesto se conoce por “inducción retrógrada”. Se aplica a menudo en teoría de juegos. La inducción retrógrada pronostica que cada jugador va a pedir 2 euros y que acabará obteniendo justamente eso: 2 euros (explicaría que el encargado de las reclamaciones en la compañía aérea haya llegado tan arriba en la empresa). Prácticamente todos los modelos utilizados por los teóricos pronostican este resultado para el DV, y ambos jugadores obtendrían 98 euros menos que si cada uno de ellos optase “ingenuamente” por reclamar 100 euros, sin pararse a considerar las posibles ventajas de seleccionar una cuantía menor.

El dilema del viajero guarda relación con otro de mayor abolengo, el famoso “dilema del prisionero”. Dos sospechosos han sido detenidos por un delito grave y están siendo interrogados por separado. Cada uno de ellos tiene la opción de delatar al otro —a cambio de lenidad— o de guardar silencio (lo que dejaría a la Fiscalía con pruebas insuficientes para lograr una condena completa *si* el otro detenido también guardase silencio). No se parece a la prueba que han de pasar nuestros viajeros; sin embargo, las recompensas del dilema del prisionero no difieren, desde el punto de vista de la pura matemática, de las de una variante del DV en la que cada viajero solamente puede reclamar 2 o 3, en lugar de cualquier entero comprendido entre 2 y 100.

En los análisis teóricos se prescinde de los aspectos anecdóticos y las historias pintorescas, para abordar la matriz de pagos de cada juego, una plantilla cuadriculada que contiene toda la información relevante sobre las posibles opciones y las recompensas que corresponderían a cada jugador. Luci ha de elegir una fila de esa matriz; Pablo, una columna. Los dos números contenidos en la casilla así determinada especifican sus respectivas recompensas, sus “pagos”.

A pesar de lo que digan sus nombres, ni el dilema del prisionero ni la versión del dilema del viajero con sólo dos opciones les plantean a los jugadores un auténtico dilema. Cada participante ve una elección inequívocamente correcta, a saber, elegir 2 (que en el caso del Prisionero consiste en delatar al otro). Tal elección se denomina “opción dominante”, porque es la mejor disponible para cada jugador con independencia de lo que haga el otro. Si Luci elige 2 en lugar de 3, recibirá 4 euros en lugar de 3 euros si Pablo opta por pedir 3 euros, mientras que recibirá 2 euros en lugar de quedarse sin nada si Pablo elige 2 euros. En contraste, la versión completa del DV no tiene opción dominante. Si Pablo elige 2 o 3, Luci deberá elegir 2, pero si Pablo optase por cualquier número entre 4 y 100, Luci saldría ganando si eligiese un número mayor que 2.


John Nash

La teoría de juegos pronostica que si el dilema del viajero se juega racionalmente, se producirá el equilibrio de Nash.

Al estudiar una matriz de pagos, lo más habitual es buscar el llamado “equilibrio de Nash”, en recuerdo de John E. Nash Jr., de la Universidad de Princeton. Un equilibrio de Nash es una elección por parte de cada jugador tal, que ninguno de ellos puede mejorar su pago cambiando unilateralmente la opción que ha escogido. Fijémonos en el resultado (100, 100) del DV (la primera coordenada es la elección de Luci, la segunda, la de Pablo). Si Luci modificase su elección a 99, pasando a (99, 100), su ganancia sería de 101 euros. Dado que Luci saldría ganando con este cambio, el resultado (100, 100) no es un equilibrio de Nash.

El juego DV tiene sólo un equilibrio de Nash —el punto (2, 2)— en el que Luci y Pablo optan ambos por 2. Lo muy difundido que está el uso del equilibrio de Nash es la causa de que tantos análisis formales pronostiquen este resultado para DV.

Existen en teoría de juegos otras nociones de equilibrio, como el equilibrio estricto, la solución racionalizable, el equilibrio perfecto, el equilibrio fuerte y varios más. Pero cada una de ellas desemboca en el pronóstico (2, 2) para el dilema del viajero. Y ahí reside el problema. La mayoría de nosotros, en nuestro fuero interno, adoptaríamos, por término medio, un valor mucho mayor que 2 euros. Nuestra intuición parece contradecir a la teoría de juegos.

Resumen/*Irracionalidad razonable*

- En el juego llamado “dilema del viajero” dos personas eligen por separado un número entero comprendido entre 2 y 100. El jugador cuyo número sea menor es recompensado con una cantidad adicional. La teoría de juegos, por consideraciones de lógica, pronostica que los jugadores deberían elegir el valor 2, pero casi todo el mundo opta por un valor más cercano a 100.
- Es necesario un nuevo tipo de razonamiento para lograr una comprensión rigurosa de que la elección racional antedicha no es racional.
- Los resultados del dilema del viajero contradicen que la teoría de juegos estándar pueda pronosticar cuál será el comportamiento de unos hipotéticos individuos egoístas y racionales, como se presume en economía. Y al mismo tiempo, hacen ver que el egoísmo no siempre constituye la mejor conducta económica.

Consecuencias para la economía

El juego, y nuestra predicción intuitiva de su resultado, está en contradicción con las ideas de los economistas. La economía, en sus albores, estaba firmemente asentada en una presunción del liberalismo, a saber, que los in-

BOB MATTHEWS *Oficina de Comunicación, Universidad de Princeton*

LA MATRIZ DE PAGOS EN EL DILEMA DEL VIAJERO

En la matriz de pagos se resume todo cuanto se necesita saber para analizar el dilema del viajero. Las opciones que Luci puede ejercitar aparecen en la columna de la izquierda (*ocre*), mientras que las de Pablo figuran en la primera fila (*amarillo*). El primer número de la casilla situada en la intersección de la fila y la columna elegidas es la remuneración (el "pago") de Luci; el segundo, el pago de Pablo. Si, por ejemplo, Luci elige 98 y Pablo toma 99, resultará que Luci recibirá 100 euros y Pablo 96.

Que ambos jugadores elijan 2, recibiendo cada uno 2 euros (*dorado*), es el equilibrio de Nash: Luci sale perdiendo (pues ganaría 0 euros) si eligiese cualquier otro número y Pablo persistiera en el 2. Análogamente, Pablo saldría perdiendo si eligiese cualquier valor distinto de 2.

Cuando las opciones se restringen a 2 y 3 (*recuadro negro*), el juego resulta equivalente al dilema del prisionero.

OPCIONES DE PABLO (euros)

OPCIONES DE LUCI (euros)

	2	3	4	...	98	99	100
2	2 2	4 0	4 0	...	4 0	4 0	4 0
3	0 4	3 3	5 1	...	5 1	5 1	5 1
4	0 4	1 5	4 4	...	6 2	6 2	6 2
...	...	...	...		...	...	...
98	0 4	1 5	2 6	...	98 98	100 96	100 96
99	0 4	1 5	2 6	...	96 100	99 99	101 97
100	0 4	1 5	2 6	...	96 100	97 101	100 100

dividuos deberían ser dejados a su albedrío, porque sus decisiones, fundadas en el egoísmo, darían por resultado un eficiente funcionamiento del sistema económico. El auge de los métodos de la teoría de juegos ha hecho ya mucho para que la economía corte amarras con esta presunción. Sin embargo, esos métodos han estado largo tiempo fundados en el axioma de que los individuos van a optar por decisiones racionales de carácter egoísta que la teoría de juegos puede pronosticar. El DV viene a minar tanto la creencia del liberalismo en que el egoísmo irrestricto confiere eficiencia al sistema económico como el principio de la teoría de juegos de que los individuos serán egoístas y racionales.

En el DV, el resultado "eficiente" es que ambos jugadores reclamen 100, porque ese resultado ofrece la máxima ganancia total a los dos viajeros. El egoísmo liberalista provocaría que ambos se apartasen de 100 hacia valores menores, con una eficiencia menor, basándose en la esperanza de una mayor ganancia individual.

Pero si los individuos no adoptan el equilibrio de Nash (2), las hipótesis en que se basa la teoría económica de comportamiento racional deben ser revisadas. Evidentemente, DV no es el único juego que pone en tela de juicio la convicción de que los individuos toman siempre decisiones racionales egoístas [*véase* "La economía del juego limpio", por Karl Sigmund, Ernst Fehr y Martin A. Nowak; INVESTIGACIÓN Y CIENCIA, marzo de 2002]. Pero sí hace ver algo más desconcertante: aunque los jugadores no se preocupen más que de su propio beneficio, no es racional que jueguen de la forma pronosticada por el análisis formal.

El dilema del viajero tiene otras consecuencias para nuestra forma de entender las situaciones del mundo real. Nos hace ver la carrera de armamentos como un proceso gradual, que nos lleva, paso a paso, hacia resultados cada vez peores. Los teóricos han tratado asimismo de generalizar el DV para tratar de comprender cómo dos empresas en competencia pueden ir cada una rebajando los precios de la otra en detrimento propio (aunque, en este caso, para beneficio del consumidor que adquiera sus productos).

Todas las consideraciones anteriores conducen a dos cuestiones: ¿Cómo actúan realmente los individuos en este juego? Y si la gran mayoría de los jugadores opta por un número mucho mayor que 2, ¿será posible explicar por qué la teoría de juegos no acierta a pronosticar que así va a ser? Sabemos ya mucho acerca de la primera cuestión; de la segunda, casi nada.

Comportamiento en la realidad

Se han realizado muchos experimentos con el dilema DV a lo largo de los diez últimos años. Han elucidado algo la cuestión. Hace años, en la Universidad de Virginia, C. Monica Capra, Jacob K. Goeree, Rosario Gómez y Charles A. Holt llevaron a cabo un experimento hoy famoso. Los jugadores eran estudiantes de ciencias económicas, y los pagos, dinero de curso legal. A los estudiantes se les primaba con 6 dólares por participar, y podían quedarse con el dinero que ganasen. Para mantener el presupuesto en límites manejables, las opciones reportaban centavos, no dólares. El intervalo de opciones iba de 80 a 200, y los valores de la recompensa y la penalización se varió en las distintas repeticiones, desde 5 hasta 80 centavos. Los experimentadores querían saber si la modificación de los importes del premio y del castigo influiría en el modo de jugar. Esas modificaciones

MATT COLLINS

no cambian en nada el análisis formal: la inducción retrógrada lleva siempre al resultado (80, 80), que es el equilibrio de Nash en todos los casos.

El experimento confirmó lo que intuitivamente se esperaba, a saber, que el jugador medio no opta por la estrategia de equilibrio de Nash, que es el valor 80. Con una recompensa de 5 centavos, la media de los valores elegidos por los jugadores fue de 180, pero descendió a 120 cuando el premio se elevó a 80 centavos.

Capra y sus colegas estudiaron también cómo se modificaba el comportamiento de los participantes al jugar repetidamente el DV. ¿Aprenderían a optar por el equilibrio de Nash, aunque no hubiera sido ésa su primera intención? Desde luego, cuando la recompensa era grande, el juego, en las sucesivas repeticiones, fue convergiendo hacia el equilibrio de Nash, el valor 80. Curiosamente, sin embargo, en el caso de recompensas pequeñas, el juego evolucionaba hacia el extremo opuesto, el valor 200.

Esta observación, a saber, que pocos optan por jugar el equilibrio de Nash, recibió confirmación adicional en un experimento acometido en la Red, aunque sin pagos en efectivo. Lo realizó Ariel Rubinstein, de las universidades de Tel Aviv y Nueva York, entre 2002 y 2004. En el juego se les pedía a los participantes, que iban a asistir a alguna de las lecciones de Rubinstein sobre teoría de juegos y Nash, que eligiesen un entero comprendido entre 180 y 300, entendido como un importe en dólares. La recompensa y la penalización se fijaron, ambas, en 5 dólares.

Participaron unas 2500 personas de siete países, lo que proporcionaba una variedad y tamaño de la muestra que no se habría podido conseguir en una prueba de laboratorio. Menos de uno de cada siete de esos jugadores optó por el equilibrio de Nash, cuyo valor era 180. La mayoría (55 por ciento) se inclinó por el valor máximo, 300. Los datos, sorprendentemente, fueron muy parecidos para los diferentes subgrupos, así como para personas de diferentes países.

Sin embargo, los procesos mentales que producen esta pauta de elecciones siguen siendo misteriosos. En particular, la respuesta más popular —la moda de la encuesta— fue 300, que es la única estrategia del juego que es "dominada", lo que significa que existe otra estrategia (299) que nunca obtiene peor resultado que aquélla y a veces la mejora.

Rubinstein dividió las posibles elecciones en cuatro grupos de números, y conjeturó que a cada uno de ellos subyace un proceso cognitivo diferente: 300 es una respuesta emocional y espontánea. La elección de un número comprendido entre 295 y 299 supone razonamientos de estrategia (por ejemplo, cierto grado de inducción retrógrada). Cualquier valor entre 181 y 294 parece tener mucho de elección al azar. Y por fin, la teoría de juegos estándar explica la elección 180, aunque los jugadores pudieron haberlo razonado por sí mismos o tener conocimientos previos sobre el punto de equilibrio del juego.

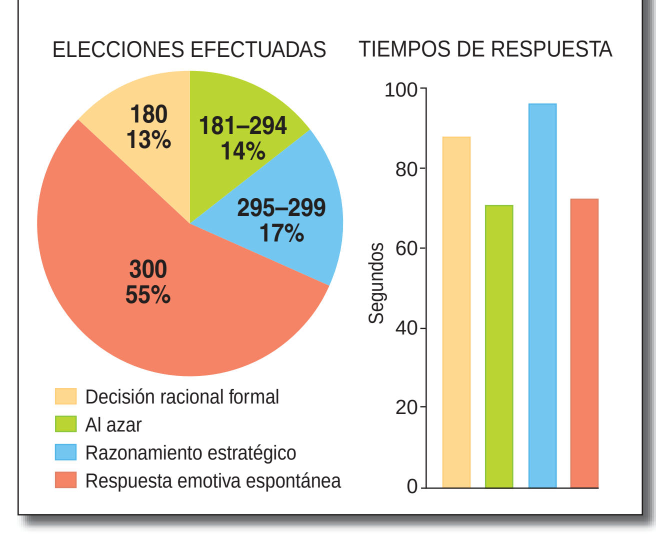
Una forma de poner a prueba la conjetura de Rubinstein para los tres primeros grupos consiste en ver cuánto tardó cada jugador en llegar a su decisión. Los que eligieron entre 295 y 299 fueron quienes, en promedio, más se lo pensaron (96 segundos), mientras que tanto los intervalos 181-294 como el valor 300 necesitaron unos 70 segundos. Esta distribución concuerda con su conjetura de que quienes eligieron entre 295 y 299 se lo pensaron más que los otros.

Los especialistas en teoría de juegos han adelantado diversas propuestas para explicar los motivos de que sean tantos los jugadores que no optan por el equilibrio de Nash en los experimentos DV. Ciertos analistas han aducido que muchas personas son incapaces de efectuar el razonamiento deductivo requerido y toman, por consiguiente, decisiones irracionales, poco meditadas. Esta explicación puede ser acertada en algunos casos, pero no para la totalidad de los resultados.

No lo es, por ejemplo, para los obtenidos en 2002 por Tilman Becker, Michael Carter y Jörg Naeve, entonces de la Universidad de Hohenheim. En su experimento, 51 miembros de la Asociación de Teoría de Juegos, profesionales de la disciplina, abordaron la versión original del dilema del viajero, cuyos extremos son 2 y 100. Cada uno de ellos jugaba contra cada uno de sus 50 oponentes seleccionando una estrategia que había de enviar a los investigadores. Tal estrategia podía consistir en la elección de un número para todas las partidas, o bien de una selección de números, acompañada de la frecuencia con la que habría de ser utilizado cada uno. La remuneración era en dinero contante y sonante: los experimentadores seleccionarían al azar a uno de los participantes, al que abonarían 20 dólares multiplicados por la ganancia media que hubiera obtenido en las 50 parti-

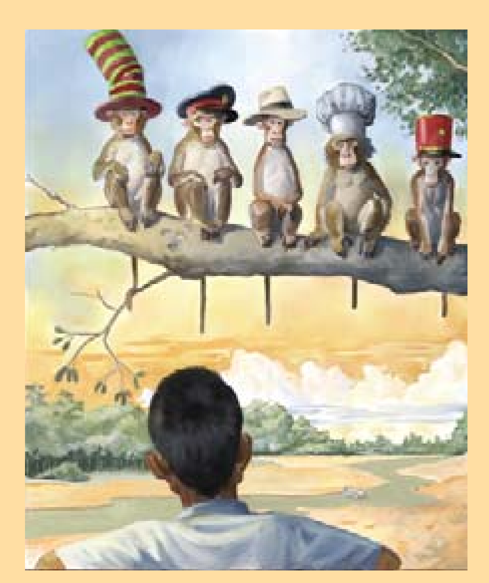
¿QUE PUDIERON PENSAR?

Se ha conjeturado que las decisiones de los jugadores en una versión del dilema del viajero donde las opciones iban desde 180 a 300 correspondían a procesos mentales distintos (*véase el diagrama sectorial*): una respuesta emocional espontánea (*la elección 300*), una elección razonada estratégicamente (*295-299*) y una elección al azar (*181-294*). Los jugadores que se atuvieron a la racionalidad formal (*180*) podrían haberla deducido por sí mismos o conocerla de antemano. Como cabría esperar, quienes eligieron de forma "espontánea" o "al azar" tardaron menos en tomar la decisión (*diagrama de barras*).



JEN CHRISTIANSEN

Teoría de juegos frente a teoría ordinaria de la decisión: yo sé que tú sabes que yo sé...



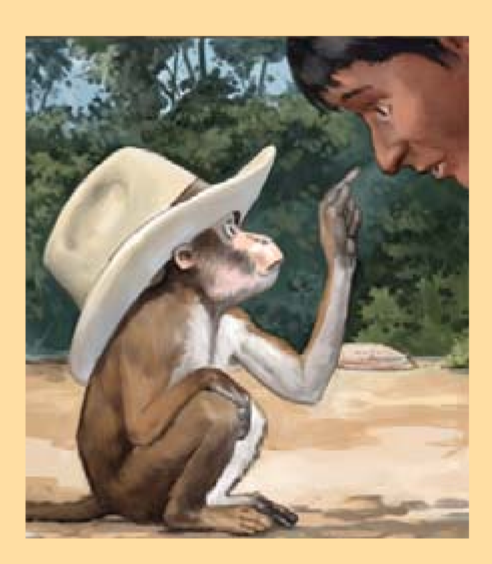
He oído este cuento en la India. Un vendedor de sombreros, al despertarse tras echar una cabezada a la sombra de un árbol, se encontró con que unos monos habían cogido todos sus sombreros y los habían subido a lo alto del árbol. Enrabietado, se quitó el sombrero que llevaba puesto y lo tiró al suelo. Los monos, que tanta tendencia tienen a la imitación, también arrojaron al suelo los suyos, que el vendedor se apresuró a recoger.

Medio siglo después, su nieto, también vendedor de sombreros, colocó su mercancía debajo del mismo árbol y se echó una siestecita. Al despertarse, vio con desesperación que los monos se habían llevado todos sus sombreros a la copa del árbol. Se acordó entonces de la historia de su abuelo y tiró con fuerza su sombrero al suelo. Pero, misteriosamente, ninguno de los monos arrojó al suelo sombrero alguno. Lo que sí ocurrió es que uno de los monos bajó del árbol, recogió el sombrero tirado en el suelo, se acercó hasta el vendedor, le dio una palmadita en la espalda y le dijo: "¿Creías que los demás no hemos tenido abuelo?"

El cuento hace ver una importante diferencia entre la teoría ordinaria de la decisión y la teoría de juegos. En esta última, lo que es racional para un jugador puede depender de lo que sea racional para el otro jugador. Para que Luci llegue a una decisión correcta, ha de ponerse en el lugar de Pablo y considerar lo que éste anda pensando, proceso que desemboca en una regresión sin fin. En teoría de juegos se describe dicha situación diciendo que la "racionalidad es de conocimiento común entre los jugadores". Con otras palabras: Luci y Pablo son racionales, cada uno de ellos sabe que el otro es racional, cada uno de ellos sabe lo que el otro sabe, y así sucesivamente.

El supuesto de que la racionalidad es de conocimiento común resulta tan habitual en la teoría de juegos, que rara vez se enuncia explícitamente. Aun así puede ocasionar problemas. En ciertos juegos que tienen un desarrollo temporal, como las repeticiones de partidas del dilema del prisionero, los jugadores pueden realizar jugadas incompatibles con ese supuesto.

En mi opinión, la presunción de que la racionalidad es de conocimiento común constituye la fuente de conflictos entre la lógica y la intuición; creo también que en el caso del dilema del viajero la intuición acierta y está en espera de una validación por una lógica más adecuada. El problema se parece a lo ocurrido en los primeros tiempos de la teoría de conjuntos. Se dio por evidente la existencia de un conjunto universal, un conjunto en el que estaría contenido todo. El conjunto universal parecía extraordinariamente natural y evidente, pero al final resultó que el origen de cierto número de paradojas de la teoría de conjuntos podía atribuirse a la suposición de que tal conjunto existía, cosa que en la matemática actual se sabe que es errónea. En mi opinión, al conocimiento común de racionalidad que se da por supuesto en teoría de juegos le espera una suerte parecida.



MATT COLLINS

das. El afortunado obtuvo finalmente 1700 dólares, por haber sido su ganancia media de 85.

De los 51 jugadores, 45 optaron por elegir un mismo valor para todas las partidas, mientras que los otros 6 especificaron más de un número. Solamente 3 de los 45 eligieron el valor 2 (el equilibrio de Nash), 10 optaron por la estrategia dominada (100), y 23 seleccionaron valores comprendidos entre 95 y 99. Es de presumir que los profesionales de la teoría de juegos sepan razonar con lógica; incluso así, fueron muy pocos, sin embargo, quienes optaron por la decisión racional que dicta la teoría.

En un análisis superficial, la explicación de sus elecciones parecería sencilla: la mayoría de los participantes consideró con justeza que sus iguales optarían sobre todo por valores comprendidos entre 95 y 100, por lo que al seleccionar ellos mismos un valor parecido obtendrían una ganancia media elevada. Ahora bien, ¿por qué esperaba casi todo el mundo que los demás optarían por un número alto?

Es posible que el altruismo y el egoísmo se encuentren firmemente arraigados en nuestro psiquismo y que nuestra conducta sea el resultado de un tira y afloja entre ambos. Sabemos que el encargado de las reclamaciones pagará el importe máximo si los dos elegimos 100. A muchos no nos parecerá bien "dejar colgado" a nuestro compañero de viaje para ganar sólo 1 dólar más, por lo que elegimos 100 a pesar de que comprendamos con claridad que lo más racional, como individuos egoístas, sería solicitar 99.

Algunos economistas, con la pretensión de ir más lejos y explicar en mayor medida las pautas de conducta observadas en este tipo de experimentos, han planteado hipótesis muy estrictas y no demasiado verosímiles. Han deducido de ahí las conductas observadas a partir de modelos complicados. No me parece que por ese camino vayamos a aprender mucho: cuanto más rebuscados se vuelven los modelos y las hipótesis, menos luz proporcionan.

Un problema irresoluto

El problema que subsiste no es, empero, la explicación de la conducta real de personas comunes al enfrentarse

al dilema del viajero. Gracias, en parte, a los experimentos, parece verosímil que el altruismo, la socialización y el razonamiento insuficiente se conjuguen para guiar la mayoría de las decisiones de los individuos. Aunque se descontasen esto tres elementos, sigo sin creer que fueran muchos quienes optasen por el valor 2. ¿Cómo podemos explicarlo, cuando, en efecto, la mayoría sigue eligiendo noventas altos aunque posea una capacidad deductiva sin tacha y aunque prescinda del altruismo y de los escrúpulos sociales para jugar implacablemente, a fin de ganar tanto dinero como sea posible? A diferencia del grueso de la teoría de juegos moderna, que hace uso intensivo de las matemáticas pero es sencilla en cuanto se dominan sus técnicas, la cuestión de marras es difícil y exige pensar de forma original.

Imaginemos que usted y yo seamos dos de estos jugadores, astutos y sin miramientos. ¿Qué se nos pasaría por la cabeza? Yo espero que usted elija un número grande, en el intervalo de 90 a 99. En tal caso, yo no debería optar por 99, porque cualquiera que sea el número que seleccione usted en ese rango, la elección 98 sería para mí igual de buena o mejor. Pero si usted se basa en su conocimiento de la despiadada conducta humana, y razona con la misma lógica que yo, también usted eliminaría la opción 99, y por el mismo tipo de razonamiento recursivo que hubiera llevado a Luci y a Pablo a que eligiesen 2, nosotros eliminaríamos rápidamente todos los números 90 a 99. Lo que estamos viendo es que no es posible definir con precisión el "conjunto de números grandes que podría elegir lógicamente la gente sin miramientos"; hemos entrado en el terreno fragoso de tratar de aplicar la razón a premisas mal definidas.

Si yo hubiera de jugar, me diría: "Dejemos de lado la lógica de la teoría de juegos. Voy a tomar un número grande (95, posiblemente) porque sé que mi oponente va a decidirse por un valor similar, y que ambos optaremos por prescindir del argumento racional de que el número inmediatamente menor daría mejor resultado que cualquiera que haya sido elegido".

Lo interesante es que este rechazo de la lógica y la racionalidad formal entraña una especie de metarracionalidad. Si ambos jugadores discurren de esta forma metarracional, ambos obtendrán buenos resultados. La noción de conducta generada por un rechazo de la conducta racional basada en consideraciones metarracionales resulta difícil de formalizar. Pero en ella reside el paso que habrá que dar en el futuro para resolver las paradojas de la racionalidad que son la plaga de la teoría de juegos y están codificadas en el dilema del viajero.

El autor

Kaushik Basu es profesor de economía, catedrático Carl Marks de estudios internacionales y director del centro de economía analítica de la Universidad Cornell. Ha escrito numeroso artículos en revistas profesionales acerca de la economía del desarrollo, economía del bienestar, teoría de juegos y organización industrial.

Bibliografía complementaria

ON THE NONEXISTENCE OF A RATIONALITY DEFINITION FOR EXTENSIVE GAMES. Kaushik Basu en *International Journal of Game Theory*, vol. 19, págs. 33-44; 1990.

THE TRAVELER'S DILEMMA: PARADOXES OF RATIONALITY IN GAME THEORY. Kaushik Basu en *American Economic Review*, vol. 84, n.º 2, págs. 391-395; mayo, 1994.

ANOMALOUS BEHAVIOR IN A TRAVELER'S DILEMMA? C. Monica Capra et al. en *American Economic Review*, vol. 89, n.º 3, págs. 678-690; junio, 1999.

THE LOGIC OF BACKWARDS INDUCTIONS. G. Priest en *Economics and Philosophy*, vol. 16, n.º 2, págs. 267-285; 2000.

EXPERTS PLAYING THE TRAVELER'S DILEMMA. Tilman Becker et al. Documento de trabajo, Instituto de Economía, Hohenheim University, 2005.

INSTINCTIVE AND COGNITIVE REASONING. Ariel Rubinstein. Disponible en arielrubinstein.tau.ac.il/papers/Response.pdf

Lepidosaurios en movimiento: de los lagartos a las serpientes

"Marinistas" y "terrestristas" sostienen ideas diferentes acerca de la historia de las serpientes: ¿se originaron en el mar o en tierra? ¿En el Cretácico o en el Jurásico?

Sebastián Apesteguía

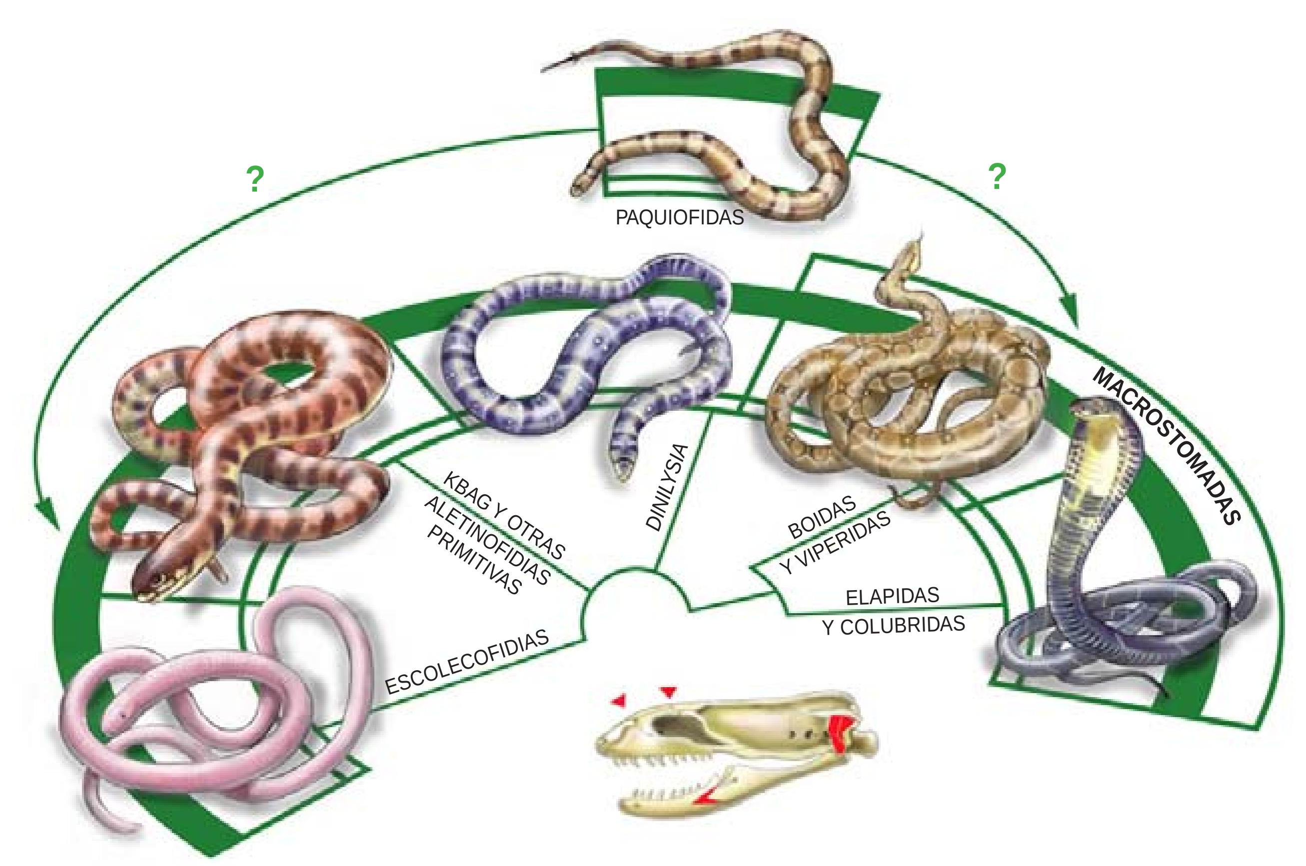
Las serpientes conforman un taxón extenso de lagartos ápodos de notable éxito evolutivo. Se ignora bastante sobre su origen. Se supone que aparecieron en Gondwana, hace al menos unos 120 millones de años. Se agrupan en dos formas principales, escolecofidias y aletinofidias, que suman unas 2700 especies. Las primeras abarcan los grupos más extraños de serpientes. Aunque se las considera primitivas, se trata de serpientes pequeñas, ciegas y subterráneas ("fosoriales"), de cráneos modificados y sólidos, con rudimentos de cintura pélvica. Sus esqueletos no arrojan mucha luz sobre el origen de las serpientes, pero han permitido suponer que las primeras fueron también fosoriales. El grupo de las aletinofidias, por otro lado, incluye las serpientes comunes, con algunas formas fósiles y todos los grupos modernos (boas, culebras, crótalos, etcétera).

Marinistas y terrestristas: el origen de las serpientes

Desde hace tiempo se viene debatiendo el origen de las serpientes. Los hallazgos de serpientes antiguas en sedimentos marinos del Cretácico Superior de Medio Oriente han abierto nuevas posibilidades de interpretación. Si estas serpientes acuáticas, las paquiófidas (*Pachyrhachis*, *Haasiophis*, *Podophis* y otras), fuesen formas primitivas o basales, el grupo no se habría originado mucho antes del Cretácico Inferior (período que se extendió entre hace unos 100 y unos 145 millones de años). Además, el hábitat marino de las paquiófidas supondría que las serpientes aparecieron en un medio acuático, lo que acentuaría la posible relación de parentesco con los mosasauroideos, grandes lagartos acuáticos, de acuerdo con la opinión de los "marinistas".

Según los "terrestristas", las paquiófidas serían serpientes de tipo moderno —es decir, aletinofidias macrostomadas, o de "boca grande"— que ni siquiera formarían un grupo con estrechas relaciones entre sí. De ser así, la gran radiación de las serpientes habría estado ya muy avanzada para el Cretácico Inferior: su origen se remontaría a algún momento del Jurásico (entre hace unos 145 y 200 millones de años). Cambiaría mucho el ambiente en que habrían aparecido; no tendría por qué ser acuático. Si los terrestristas, que consideran a las paquiófidas serpientes modernas, andan en lo cierto, las serpientes guardan el parentesco más estrecho con algún grupo de lagartos ápodos de hábitos subterráneos y no demasiado alongados, como los ánguidos, dibámidos o anfisbenas. En consecuencia, sería razonable esperar el hallazgo de serpientes terrestres basales a lo largo de todo el Cretácico de Gondwana.

La hipótesis terrestrista, o continentalista, explicaría también el sistema de locomoción terrestre de las serpientes actuales. El mismo movimiento que les sirve para moverse con facilidad en el agua, una simple ondulación lateral, les vale también en tierra firme ejerciendo fuerza contra las irregularidades del suelo. También pueden angostar o ampliar alternativamente distintas partes del cuerpo gracias a su musculatura y costillas fijas, que anclan una parte y hacen avanzar otra.



JORGE A. GONZALEZ; DISEÑO DE MARIANO SANSIBIERI

1. LAS SERPIENTES han tenido considerable éxito desde sus orígenes, a comienzos del Cretácico. Entre las más primitivas aún vivientes se hallan las escolecofidias subterráneas y las anilioideas, primitivas aletinofidias. Por su parte, las macrostomadas incluyen a casi todas las serpientes modernas (boas, víboras, culebras). Las paquiófidas marinas del Cretácico de Israel y Europa Oriental poseen a la vez características primitivas y avanzadas; por esa razón, su posición es incierta. El cráneo de una serpiente anilioidea (*abajo*) muestra las principales características del grupo: junta intramandibular móvil, hueso de articulación mandibular (*cuadrado*) colgante y articulaciones laxas en el techo del cráneo.

Pero tal vez su modo de locomoción más sorprendente sea el "caminar con las escamas", que realizan mediante movimientos alternados de series de escamas del vientre que se orientan controladas por la musculatura epidérmica. Si su origen es terrestre, y cabe presumir que fosorial, debería hallarse ligado también a la presencia de vértebras con espinas neurales bajas y tamaños corporales no demasiado grandes. ¿Existieron serpientes así?

Najash, la madre de las serpientes

Una antigua serpiente fósil de menos de 2 metros de largo y 90 millones de años de antigüedad se descubrió en La Buitrera, un yacimiento fosilífero cercano a Cerro Policía y a unos 1300 kilómetros de Buenos Aires, en la provincia de Río Negro. Realizó el hallazgo el equipo de exploración del Museo Argentino de Ciencias Naturales de Buenos Aires y la Fundación de Historia Natural Félix de Azara (CEBBAD-Universidad Maimónides), dirigido por el autor. Desde 1999, las expediciones a La Buitrera han aportado una impresionante lista de hallazgos novedosos: dinosaurios (*Buitreraptor*, *Catharte-saura*), reptiles esfenodontes (*Priosphenodon*, dado a conocer en 2003), mamíferos, tortugas y peces.

En la expedición de 2001 y 2002, Pablo Gallina, miembro del equipo, halló restos de serpiente. La expedición de 2004 dio con fragmentos de un cráneo. Tras varios años de preparación del fósil y estudios comparados en la Argentina, Brasil y Estados Unidos, en colaboración con Hussam Zaher, del Museu de Zoologia de la Universidad de São Paulo, ha sido posible comprender la relevancia de esta serpiente fósil, a la que se ha dado el nombre de *Najash rionegrina*. Constituye la más primitiva de las serpientes conocidas. El nombre proviene de la palabra hebrea *Najash*, serpiente, y es el nombre aplicado a la serpiente en el *Génesis*. La especie, *rionegrina*, se refiere a la provincia donde se la encontró.

El hallazgo de *Najash* en el Cretácico de la Patagonia, una serpiente terrestre, con patas posteriores robustas y dientes semejantes a los de las serpientes aletinofidias, aporta una perspectiva inédita para el debate. No se trata de otra serpiente marina, sino de una serpiente que vivió en un ambiente terrestre, a la sombra de los dinosaurios y en América del Sur. Las serpientes como *Najash* eran bastante pequeñas, vivían en cuevas y se alimentaban de lagartos, micromamíferos y crías de dinosaurio. Sobresalen sus patas bien desarrolladas y ancladas en la columna vertebral por la región sacra; demuestran que les eran útiles para moverse o, tal vez, para anclarse en la tierra al atacar.

Las primeras aletinofidias

Existe una agrupación artificial de serpientes halladas a lo largo de todo el Cretácico de Gondwana, muy relacio-

nadas con las primeras aletinofidias y que guardan numerosas características en común, en particular determinados orificios y espinas neurales bajas en sus vértebras. Le concederemos la denominación provisional KBAG (K es el símbolo de Cretácico y el resto significa Basal, Aletinofidia y Gondwana). Se ha repartido estas serpientes en dos grupos: laparentófidas y madtsoidas. Este último nombre, que no compartimos, da por sentado una cercana relación con la enorme serpiente *Madtsoia bai*, del Eoceno de Argentina, aunque no hay base observacional suficiente.

Entre los restos de KBAG de mayor antigüedad se hallan las vértebras aisladas de *Lapparentophis defrennei*, del Cretácico Inferior del Sahara, de unos 110 millones de años de antigüedad, así como vértebras del Cretácico Temprano de España, algo más viejas; más dudosas también. En Argentina, el registro de estas serpientes incluye materiales sueltos abundantes de finales del Cretácico del litoral norpatagónico. ¿Fueron, pues, las KBAG serpientes aletinofidias primitivas?

Ligada a las aletinofidias primitivas se encuentra la serpiente *Dinilysia patagonica* (según algunos, está muy cerca de ellas pero no pertenece al grupo; según otros, sí sería una aletinofidia). Desde principios del siglo XX se vienen desenterrando huesos de este animal, muy completos, en rocas del Cretácico Superior de las cercanías de la ciudad de Neuquén, en el norte patagónico. Tales fósiles comprenden varios esqueletos con cráneo, vértebras y costillas articuladas; no se han hallado miembros posteriores, pero es posible que los tuviera.

Si bien se creyó en un principio que *Dinilysia* estaba emparentada con las boas modernas, su estudio detenido descubre que se trata de un parecido superficial. Sí comparte numerosas características con las KBAG. Lo mismo éstas que *Dinilysia* se hallan cercanas al origen de las serpientes aletinofidias. El linaje al que *Dinilysia* pertenece retuvo algunas características primitivas.

¿Cómo siguió la evolución de las KBAG? Algunos piensan que las madtsoidas genuinas, serpientes gigantes del Eoceno de Patagonia, son miembros de gran tamaño del mismo linaje y que habrían vivido incluso hasta el Cuaternario de Australia, como *Wonambi*. Para otros, las madtsoidas terciarias son serpientes modernas macrostomadas, tal vez emparentadas con las boas, sin relación, o muy escasa, con las KBAG, excepto por haber retenido ciertas características primitivas que les otorgan una semejanza superficial. Con ese mismo grupo habría estado emparentada *Gigantophis*, del Terciario de Egipto.

Ya que las serpientes escolecofidias están demasiado modificadas por su vida subterránea, las boas son un grupo de serpientes modernas y de las paquiófidas piensan algunos que pertenecen también a un grupo de serpientes modernas que se habría dispersado pronto por el mar de Tethys, entre los continentes Laurasia y Gondwana, ¿hay acaso algún grupo que nos permita interpretar cómo fueron las primeras serpientes?

Del grupo cretácico de aletinofidias basales hay aún algunos representantes vivientes, en especial las anilioideas (por ejemplo *Anilius*, *Cylindrophis*, *Uropeltis* y *Anomochilus*). Si bien son plenamente serpientes, su morfología craneana nos permite compararlas con diversos linajes de lagartos. Provenientes de Sudamérica y el sudeste asiático, estas serpientes se alimentan de gusanos, peces, anfibios y de escolecofidias. A través de ellas podemos sospechar cómo se alimentaban las serpientes



JORGE A. GONZALEZ

2. LOS BUITRERRAPTORES, pequeños dinosaurios carnívoros parientes del *Velociraptor*, seguramente se alimentaban también de serpientes primitivas como *Najash*; las extraerían de sus cuevas con sus alargados hocicos.

primitivas, capaces sólo de expandir sus bocas moderadamente, mucho menos que la mayoría de los ofidios de hoy, pero más que un lagarto. Estas aletinofidias basales vivientes, o anilioideas, a diferencia de las serpientes modernas, poseen un hueso maxilar poco móvil (aunque más que el de un lagarto) y la región del hocico articulada con la angosta región posterior mediante una articulación entre los huesos nasal y frontal. El hueso cuadrado, donde se articula la mandíbula, es corto y vertical en estas serpientes.

Macrostomadas

En las serpientes modernas, distingue a las macrostomadas un conjunto de características combinadas que les confieren una notable movilidad craneana. Al moverse los huesos que componen sus cráneos y mandíbulas en tres dimensiones, sus bocas se expanden de un modo desmesurado ("macrostomadas" significa exactamente eso, "boca grande").

Entre las características aludidas merecen citarse la pérdida de la barra temporal superior (además de la inferior, que ya estaba ausente originalmente en los lepidosaurios), una muy flexible conexión entre cráneo y mandíbulas, una suave unión entre ambas hemimandíbulas, y entre la mitad anterior y posterior de cada una de ellas, en la "junta intramandibular". (Si bien desde 1870 se suponía que tanto serpientes como mosasauroideos compartían esta unión intramandibular, recientes investigaciones han demostrado que se habrían desarrollado en cada grupo por su lado.)

La apertura bucal de las macrostomadas les permite ingerir presas mayores que sus propias cabezas; lo logran gracias a los movimientos, independientes de cada maxilar y de cada hueso del paladar, que efectúan la "caminata pterigoide": esos huesos dentados avanzan de manera independiente sobre la presa, mientras las mandíbulas la van acomodando por debajo y los labios la abarcan.

Para algunos, las semejanzas en la estructura craneana de los varanos y las serpientes son ineludibles. El "monitor sin oídos" (*Lanthanotus*), un lagarto varanoideo semiacuático de Indonesia, presenta además otras características semejantes a las de las serpientes, como una cubierta escamosa modificada que protege del agua y la suciedad a los ojos, ausencia de oídos y una importante reducción en los miembros, extrema en las serpientes.

Otro carácter que relaciona a éstas con los varanoideos es la lengua bífida, que toma sustancias químicas volátiles del aire y las acerca a dos aberturas que tienen en el techo de la boca ligadas al órgano de Jacobson. Situado éste en la parte anterior del paladar, se trata de un tejido semejante al olfatorio, aunque plegado para multiplicar la eficiencia de la detección. Este órgano es frecuente entre los vertebrados, pero en las serpientes está muy desarrollado: pueden seguir un rastro químico específico durante largas distancias, y las "texturas", "intensidades" y "matices" que captan les ofrecen un panorama casi tan complejo como a nosotros el mundo visual. Por otro lado, la dentina plegada hacia el interior del diente que presentan los varanos (plicidentina), y que supuestamente caracterizaba a ambos grupos, no resultó ser tal en las serpientes ni en los mosasaurios, sino sólo surcos en el esmalte, comunes en muchos vertebrados.

Si el escamado que originó a las serpientes se asemejó al monitor sin oídos, las serpientes derivarían de un ancestro semiacuático, como este lagarto piscívoro. Esa conclusión sustentaría la teoría de los marinis-



SEBASTIAN APESTEGUIA

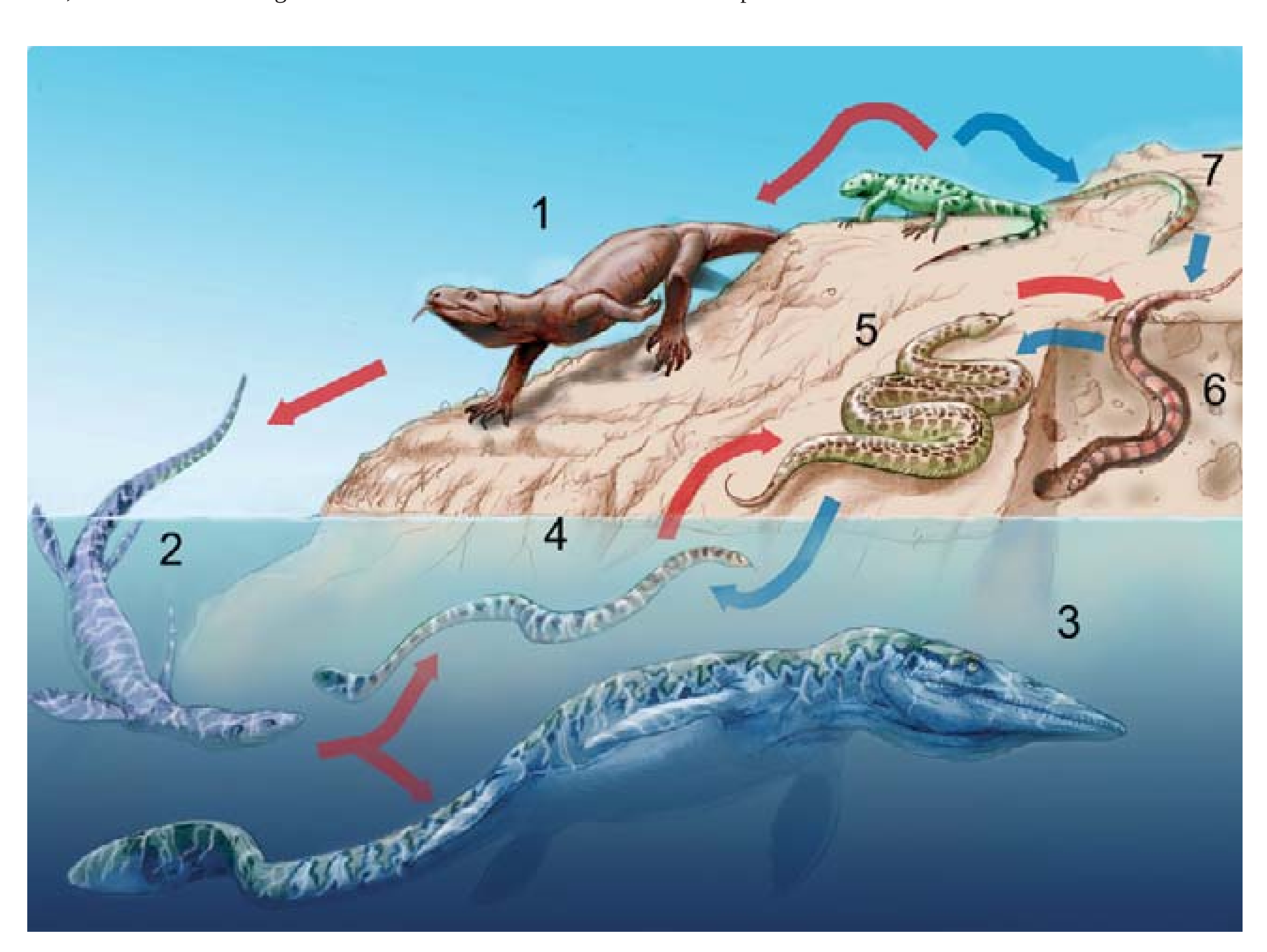
3. NAJASH RIONEGRINA es una primitiva serpiente con patas posteriores que vivió en la Patagonia argentina hace 92 millones de años, a la sombra de los más gigantescos dinosaurios, como *Argentinosaurus* y *Giganotosaurus*. La foto muestra detalles de su anatomía.

tas, avalada por las paquiófidas. Sin embargo, muchas de las adaptaciones que remiten a un origen marino, en especial la pérdida de miembros y el alargamiento del tronco, concuerdan también con un modo de vida fosorial. Además, el fortalecimiento del cráneo, que incluye el cierre y protección del cerebro con paredes óseas (en los lagartos hay sólo tejidos blandos), junto con la adquisición de tamaños muy pequeños (una miniaturización), inclinan a pensar en un modo de vida en el que el lagarto buscara a sus presas sobre el suelo (esas mismas adaptaciones resultan frecuentes en los lagartos ápodos, por ejemplo en las anfisbenas).

La adopción de un modo de vida subterráneo, a diferencia del acuático, suele llevar consigo la atrofia de la audición y la visión, con todos sus componentes asociados. Las vibraciones del suelo les llegan a las mandíbulas; desde allí, se transmiten a través del hueso cuadrado, la columela (nuestro "estribo") y el oído interno. Las serpientes no subterráneas ven, pero la región de su cerebro relacionada con la visión y la estructura misma de sus ojos, incluido el mecanismo de enfoque, son tan diferentes de los presentes en los lagartos, que parecen haberse originado de nuevo a partir de ancestros no videntes o de visión reducida, tal vez subterráneos. Complementan su percepción del mundo receptores térmicos y sensaciones táctiles muy desarrolladas.

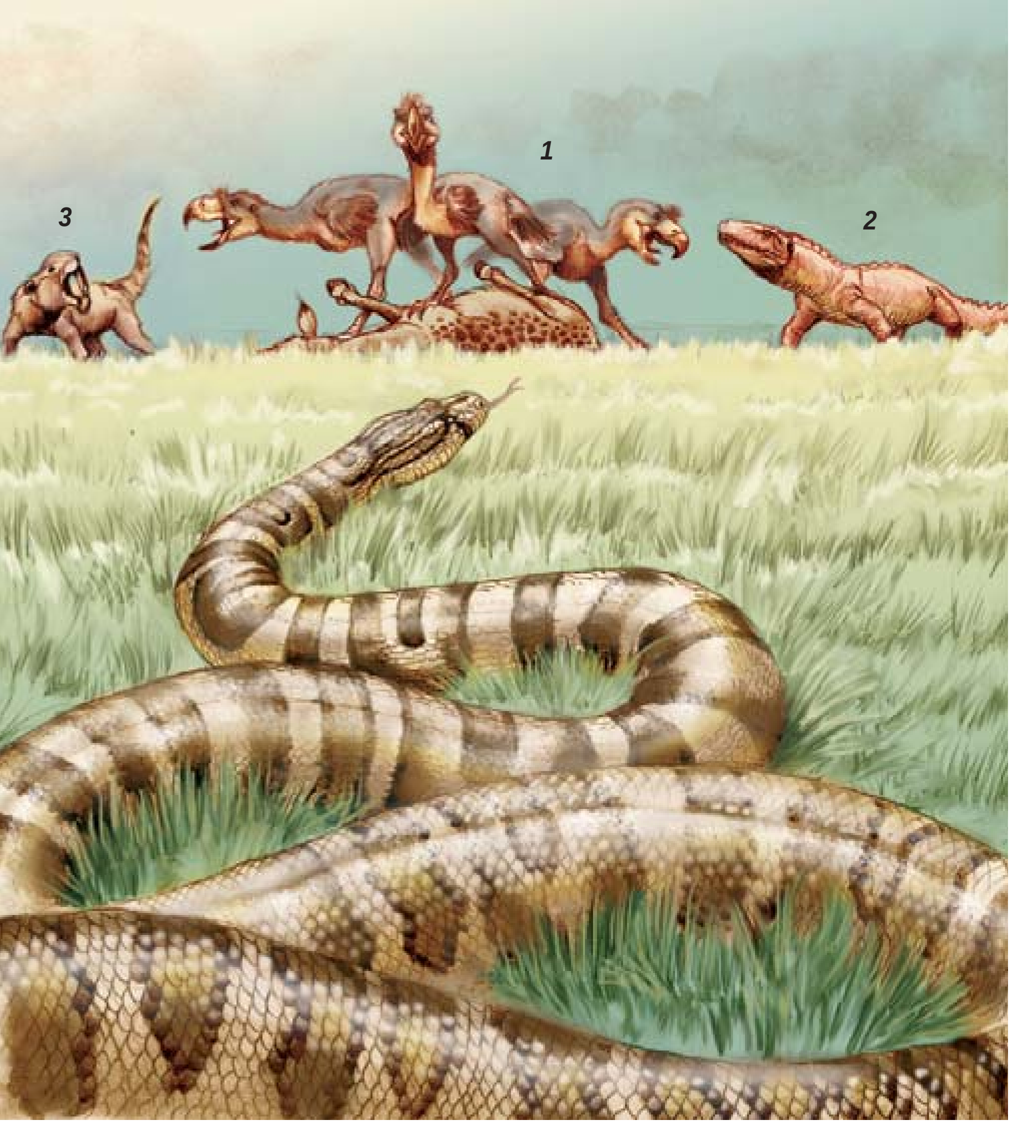
Muchos investigadores han visto en los rudimentos de huesos pélvicos y miembros una señal llamativa de primitivismo. Cierto es que la pérdida de los miembros anteriores ocurrió pronto en la evolución de las serpientes; no obstante, los miembros posteriores aún se encuentran presentes en todas las escolecofidias, las aletinofidias basales e incluso en macrostomadas primitivas, como *Acrochordus*, pariente próxima de las culebras, boas y pitones: sólo los han perdido de verdad serpientes muy avanzadas.

Dinilysia, del Cretácico Superior de la Patagonia argentina, cercana a las aletinofidias basales, muestra en su cráneo caracteres que se han relacionado con un modo de vida semiacuático. *Dinilysia*, sin embargo, es un miembro tardío dentro de la radiación cretácica. Su tamaño



4. LA EVOLUCION DE LAS PRIMERAS SERPIENTES puede explicarse mediante dos hipótesis: la marinista, en la que lagartos relacionados a los varanoideos (*1*) que se hicieron marinos, como los eguialosáuridos (*2*), habrían dado origen tanto a los grandes mosasaurios (*3*) como a las serpientes paquiófidas (*4*) y finalmente, retornando a la tierra, a las serpientes modernas, lo mismo las de superficie (*5*) que las subterráneas (*6*). Según la hipótesis de los terrestristas, las serpientes se habrían originado a partir de lagartos ápodos (*7*), cazadores subterráneos como las actuales anfisbenas, por lo que las serpientes subterráneas (*6*) se hallarían entre las más primitivas y las de superficie (*5*) entre las derivadas de ellas.

JORGE A. GONZALEZ



5. SERPIENTES GIGANTES. Conocidas con el nombre de madtsoidas, desempeñaron el papel de grandes carnívoros terrestres en los ecosistemas del Terciario de América del Sur, en disputa con las aves fororrácidas (*1*), los cocodrilos terrestres sebécidos (*2*) y los marsupiales carnívoros (*3*). No habiendo en América del Sur ni cánidos ni félidos, el papel de carnívoros les tocó a los mismos actores que lo desempeñaron durante el Cretácico: dinosaurios bípedos (aves), cocodrilos terrestres, lepidosaurios y algunos mamíferos. De los restos del festín se encargaría luego la descomunal ave *Argentavis*.

JORGE A. GONZALEZ

poderoso y hábitat continental sugieren que pudo ser una forma de KBAG adaptada secundariamente a los ambientes lagunares. Tal hipótesis guarda coherencia lógica, si consideramos los 30 millones de años de evolución que la separan de *Najash* y una amplia dispersión por el norte y sur de Gondwana.

Dispersión

Tras una importante expansión gondwánica a lo largo del Cretácico, las serpientes cruzaron el mar de Tethys. Con mayor precisión, el mar se fue cerrando, mientras Africa, España, Turquía y otras regiones se acercaban a Europa. Hacia fines del Cretácico, las serpientes empezaron a difundirse por los ecosistemas laurásicos europeos y norteamericanos, aunque a otros lugares, como Asia Central, tardaron más en llegar.

Para el Terciario Temprano, las serpientes se habían extendido por los Trópicos. Iban ya apareciendo varios grupos modernos, entre ellos las boas genuinas. Estas macrostomadas, que pueden sobrepasar los diez metros de longitud, conservan rudimentos de la cintura pélvica y los miembros posteriores, que los machos utilizan para sujetar a la hembra durante la cópula. A diferencia de otras serpientes avanzadas, poseen premaxilares con dientes. Muchas boas tienen a los lados del rostro numerosos receptores térmicos, las “fosetas loreales”, que detectan el calor de las presas.

Presumiblemente durante el Terciario Temprano, aparecieron, en algunas serpientes, especializaciones para inocular veneno, producido por sus glándulas salivales modificadas. Esos venenos actúan de modo dispar sobre el organismo. Podemos clasificarlos, a grandes rasgos, en neurotóxicos y hemolíticos. Los primeros afectan al sistema nervioso; causan parálisis respiratoria y ceguera. Los hemolíticos destruyen irreversiblemente las células y los tejidos, con las consiguientes disfunciones y fuertes hemorragias.

En algunas serpientes, el veneno se desliza por una acanaladura abierta en los dientes de la parte posterior del maxilar (opistoglifos). En otras, muy ponzoñosas, los dientes inoculadores, que forman un surco cerrado, se hallan en la parte anterior del maxilar (proteroglifos). Y aún en otras, el maxilar es corto y móvil, de modo que el diente inoculador, largo y con el canal cerrado, convertido en una aguja hipodérmica de gran extensión (solenoglifos), se guarda en posición horizontal en la boca cerrada paralelo al paladar; al abrir la boca, los colmillos inoculadores se tornan verticales y en posición de ataque.

A este último tipo de serpientes venenosas pertenecen las víboras, que carecen de una representación fósil esclarecedora. El grupo, que no se extiende más allá del Terciario Superior, incluye también los crótalos. Poseen colmillos desproporcionadamente grandes y solenoglifos, y evolucionaron pronto, si se compara con las boas. Algunas víboras presentan una especial sensibilidad en sus dos únicos receptores térmicos loreales que les permite cazar mamíferos por la noche o seguirlos a sus cuevas subterráneas. En los crótalos, la sensibilidad de las fosetas termorreceptoras es tan alta, que captan cambios de milésimas de grado centígrado.

Algunos crótalos han desarrollado el “cascabel”, una estructura derivada de la piel que se encuentra al final de la cola y funciona como dispositivo de advertencia. Las boas y las víboras se han convertido así en

expertas cazadoras de mamíferos; las primeras utilizan la constricción y sofocación, las segundas recurren a eficaces venenos.

Hacia mediados del Terciario, a partir de serpientes de aspecto semejante a las boas, por ejemplo *Acrochordus*, se desarrollaron explosivamente nuevos grupos de ofidios de pequeño tamaño, las culebras o colúbridas. Terminaron por constituir las dos terceras partes de las serpientes actuales. Las boas se asocian hoy a los trópicos; las culebras, en cambio, presentan una mayor flexibilidad ambiental: toleran en sus hábitats rangos de temperatura más amplios, como los que imperan desde mediados del Terciario. Las culebras poseen cortos colmillos en la parte posterior de la boca y su veneno, débil, sirve de sedante para inmovilizar a la presa. Algunas culebras que se alimentan de anuros, como *Xenodon*, utilizan los dientes, muy grandes, para perforar los pulmones de sus presas, que acostumbran inflarse para aparentar un tamaño mayor y a la vez para que cueste tragarlas.

Tras el Mioceno, algunas serpientes emparentadas con las culebras desarrollaron poderosos venenos, aunque con modos de inoculación distintos de los que utilizan las víboras. Pese a que su veneno, también derivado de las glándulas salivales, es muy tóxico, los dientes inoculadores se sitúan en la parte posterior de la mandíbula. En las elápidas, grupo que incluye a las cobras, las mambas, las corales y las peligrosas serpientes de mar, los dientes, no retráctiles, han migrado hacia la región anterior de la boca, donde pueden servir de defensa. En las atractaspídidas, posibles elápidas primitivas con especies que se alimentan de anfibios ápodos (las cecilias), se han desarrollado dientes de tipo solenoglifo, como los de las víboras.

Una historia compartida

La historia de las serpientes se halla íntimamente ligada a la de los mamíferos. Si bien han surgido mucho después, resultan muy sugestivas sus adaptaciones térmicas, que les permiten cazar mamíferos con aterradora precisión. Cuando aparecieron a principios del Cretácico, en los territorios de Gondwana, los principales grupos de mamíferos presentes eran los driolestoideos y gondwanaterios, aunque también es posible que hubiera monotremas, como el ornitorrinco y también los armadillos primitivos.

Especialistas en presas pequeñas, aquellas primeras aletinofidias no precisaban de una apertura bucal exagerada para dar cuenta de ellas. Sin embargo, considerando el costo energético invertido en la cacería de micromamíferos, resulta explicable la ventaja de cazar una presa más voluminosa. Quienes pudieran, preferirían optar por las piezas más ambiciosas. Desconocemos si las primeras serpientes macrostomadas capaces de una apertura grande de boca buscaron presas mayores, pero suponemos que sí: pequeños dinosaurios (las de tierra firme) y grandes peces o pequeños reptiles marinos (las paquiófidas del mar de Tethys).

Lo sorprendente es que, según la hipótesis terrestrista, para mediados del Cretácico ya habría serpientes especializadas en la vida en varios ambientes. Esto supone una radiación evolutiva veloz, en la que no hubo tiempo suficiente para que los linajes se despojaran de algunas de sus características primitivas; por ejemplo, las patas. Por ello, no es ilógico esperar que casi todas las serpientes cretácicas conserven restos de extremidades posteriores.

Integradas en la radiación evolutiva de los lepidosaurios, las serpientes se hallan capacitadas para deglutir presas más voluminosas que ellas mismas y generar poderosos venenos, trepar por superficies imposibles, nadar e incluso planear entre los árboles de la selva. Sus colores pueden ser tanto apagados y miméticos como brillantes. Mientras nuestros ancestros primates saltaban entre las ramas, enormes serpientes, como la africana *Gigantophis*, buscaban su presa por el suelo, y no pocas veces las serpientes nos han atacado o dado horribles sustos, por lo que muchos humanos conservamos una repulsión ancestral hacia ellas.

Numerosas culturas han llevado su imagen a los altares con diferentes atributos. Fue una serpiente la que le robó la inmortalidad a Gilgamesh y engañó a Eva, y fueron serpientes las que promovieron el culto de Seth en Egipto, de Quetzalcóatl en Centroamérica y de las dos serpientes gigantes enemigas, Cai-Cai y Tren-Tren, en la Patagonia argentina. Son las serpientes la encarnación de un peligro: los indios hopis las usan en sus danzas y los hindúes en sus ritos. También los griegos simbolizaron con ellas el saber médico. Las serpientes son tan buenas cazadoras de micromamíferos, que en los lugares donde se las elimina aumenta la cantidad de ratas y agentes transmisores de enfermedades.

Aún hoy, muchas serpientes inofensivas, principalmente culebras, aunque también otros lagartos ápodos, pagan con sus vidas su semejanza con otras especies peligrosas. Son víctimas de nuestra ignorancia acerca de estos sorprendentes lepidosaurios.

El autor

Sebastián Apesteguía es paleontólogo del Museo Argentino de Ciencias Naturales, donde trabaja desde 1988, director del Area de Paleontología del CEBBAD (Fundación Félix de Azara-Universidad Maimónides) y titular de la cátedra de Herpetología de la Universidad CAECE. Ha descubierto varios yacimientos fosilíferos cretácicos, entre los que se halla "La Buitrera". Es autor de 42 artículos científicos y 4 libros de dinosaurios para jóvenes. Investiga sobre la evolución de los esfenodontes y de los dinosaurios saurópodos.

Bibliografía complementaria

LA "VISIÓN" INFRARROJA DE LAS SERPIENTES. Eric A. Newman y Peter H. Hartline en *Investigación y Ciencia*, págs. 88-98, mayo 1982.

THE ORIGIN OF SNAKES: LIMITS OF A SCIENTIFIC DEBATE. Olivier Rieppel y Maureen Kearney en *Biologist*, vol. 48, n.º 3, págs. 110-114; 2001.

VENENOS DE SERPIENTES. DIVERSIFICACIÓN ANCESTRAL DE PROTEÍNAS ANCESTRALES. Juan José Calvete, M.ª Paz Moreno Murciano y C. Marcinkiewicz en *Investigación y Ciencia*, págs. 33-35, abril 2003.

A CRETACEOUS TERRESTRIAL SNAKE WITH ROBUST HINDLIMBS AND A SACRUM. Sebastián Apesteguía y Hussam Zaher en *Nature*, vol. 440, págs. 1037-1040; 2006.

Telescopios de espejo líquido

Una vieja idea para captar imágenes astronómicas está renaciendo gracias a los avances técnicos

Paul Hickson

Hay en las remotas montañas suroccidentales de la Columbia Británica un caserón que, desde lejos, uno confundiría con una iglesia o con un refugio de esquiadores. Nadie diría, ni siquiera visto de cerca, que esconde uno de los mayores telescopios de Norteamérica. En noches claras, su empinado tejado se abre para que el telescopio obtenga imágenes de las estrellas y galaxias que pasan sobre él. Lo más sorprendente, sin embargo, es el corazón de ese telescopio, su espejo primario, que colecta y enfoca la luz: un disco rotante de mercurio líquido.

Los telescopios astronómicos corrientes utilizan espejos de vidrio para ese propósito. Los dos mayores, los telescopios gemelos Keck de 10 metros de diámetro, situados en la cima del Mauna Kea, en Hawai, disponen de sendos espejos primarios segmentados, cada uno de ellos compuesto por 36 segmentos hexagonales. Hubo que construir y pulir esos segmentos con una precisión de unas pocas décimas de nanómetro, para recubrirlos luego con una fina capa de aluminio o plata que les proporciona la reflectividad. Espejos así requieren también un complejo sistema de soportes para evitar que su superficie se distorsione con los cambios de temperatura o por efecto de la gravedad.

En la mayoría de los telescopios modernos de gran diámetro, un sistema de sensores y actuadores controla activamente la forma del espejo a pequeña escala para contrarrestar las distorsiones creadas por la atmósfera, procedimiento que recibe el nombre de "óptica adaptativa". Estas maravillas de la técnica resultan muy caras: del orden de 10 millones de dólares si el espejo mide seis metros de diámetro. Por extraño que parezca, se consigue una precisión comparable rotando un disco cubierto de mercurio.

Esta aseveración se basa en un principio elemental: la superficie de un líquido en equilibrio es una superficie de energía potencial constante (cualquier variación del potencial a lo largo de la superficie implicaría una fuerza que haría que el líquido fluyese). En condiciones normales, la energía potencial de un objeto es proporcional a su altura; por cuya razón las superficies de la mayoría de los líquidos son planas. Pero supongamos que ponemos a girar un líquido con velocidad angular constante alrededor de un eje vertical. La energía potencial de cualquier pequeña porción de fluido tiene ahora dos componentes: una que crece con la altura y otra que disminuye con la distancia al eje de rotación, o mejor dicho, con el cuadrado de esa distancia. Esta combinación de dependencias da a la superficie forma de paraboloide.

Por una feliz coincidencia, un paraboloide es exactamente lo que se necesita para enfocar la luz. Los rayos de luz incidentes paralelos al eje de un paraboloide se reflejan de suerte tal, que convergen todos en un punto, el foco del telescopio. Los rayos paralelos que lleguen con otra inclinación (procedentes de otra dirección) no se enfocarán con semejante perfección. Pero se puede obtener una imagen de buena calidad para un campo de visión extenso, si se añaden tres o cuatro lentes cerca del punto focal. La mayoría de los grandes telescopios astronómicos emplean espejos primarios parabólicos (o casi parabólicos) y espejos secundarios o lentes para corregir las aberraciones fuera de eje.

La idea de enfocar la luz con un líquido en rotación viene de antiguo. Ernesto Capocci, del Observatorio de Capodimonte en Nápoles, describió ya esa posibilidad en 1850, aunque nunca lo puso en práctica. La materialización de la idea se produjo en 1872, cuando Henry

1. EL AUTOR POSA ANTE EL GRAN TELESCOPIO CENITAL, construido por su equipo. La indumentaria protectora es necesaria en ciertas ocasiones porque el espejo de 6 metros de diámetro del telescopio no es de vidrio, sino de mercurio. El espejo gira a unas 7 revoluciones por minuto. El líquido adquiere así la forma de un paraboloide, que colecta y enfoca la luz en un detector montado sobre el foco.





2. EL OBSERVATORIO que cobija al Gran Telescopio Cenital se halla en una colina (a 395 metros de altura), en el Bosque de Investigaciones Malcolm Knapp, de la Universidad de la Columbia Británica.

Skey, del Observatorio Dunedin de Nueva Zelanda, construyó un espejo líquido de 35 centímetros de diámetro en su laboratorio. En 1909, Robert W. Wood, de la Universidad Johns Hopkins, construyó el primer telescopio completo de espejo líquido. El mejor modelo de Wood llevaba un espejo de 51 centímetros de diámetro. Giraba sobre un cojinete mecánico impulsado por un motor, con una correa de transmisión formada por finas tiras de goma natural. Con este telescopio, Wood resolvió el sistema estelar cuádruple ε Lyrae, que tiene componentes separadas por tan sólo 2,3 segundos de arco. Un logro impresionante: sólo le separaba un factor 10 del límite que la difracción impone a un espejo de ese tamaño.

Aun así, el telescopio de Wood no resultaba práctico. No podía librarse de las vibraciones; y el espejo experimentaba oscilaciones, aunque pequeñas, perceptibles. Además, un control impreciso de la velocidad causaba la fluctuación de su distancia focal. Dado que el eje de rotación tenía que ser vertical, el telescopio podía observar sólo una estrecha área del cielo, situada justo encima, mientras la rotación terrestre inducía un movimiento constante de las imágenes. Tales problemas explican por qué la astronomía dejó de lado los espejos líquidos durante los siguientes 73 años.

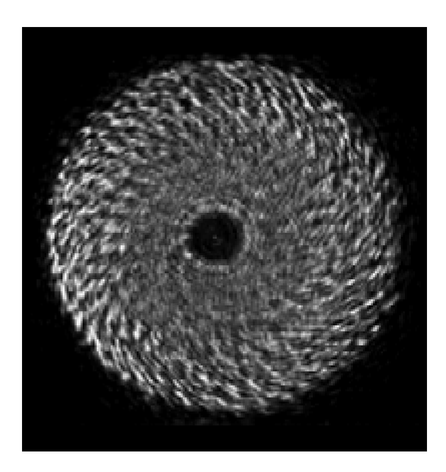
El resurgimiento

En 1982, Ermanno F. Borra, de la Universidad Laval, de Québec, decidió retomar la idea de Wood. Borra y su grupo pensaron que la técnica moderna resolvería las dificultades que Wood no pudo afrontar.

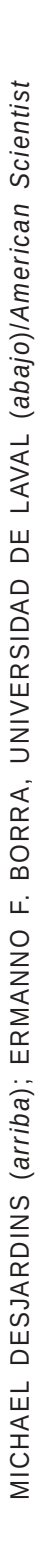
En particular, solventaron el problema de que el movimiento de la Tierra desplace la imagen a lo largo del campo de visión. Reemplazaron la película fotográfica tradicional por un detector moderno, un dispositivo de acoplamiento de carga (CCD). Estos sensores de estado sólido constituyen un tipo particular de circuito integrado de silicio. Los fotones que inciden sobre el silicio depositan energía suficiente para generar e impulsar electrones hasta la banda de conducción. Estos electrones se almacenan en pozos de potencial creados por pequeños voltajes que se aplican a una matriz de electrodos. Así, una imagen óptica se convierte dentro del silicio en una imagen electrónica. Al final de una exposición, se cierra el obturador y se manipulan los voltajes aplicados a los electrodos para que los grupos de electrones se muevan hacia un lado del detector. Desde allí salen a un amplificador de salida, que produce una serie de señales de voltaje, cada una proporcional al número de electrones atrapados inicialmente en cada pozo de potencial. En consecuencia, la imagen se lee píxel a píxel, remitidos a un ordenador que los almacena, procesa y exhibe.

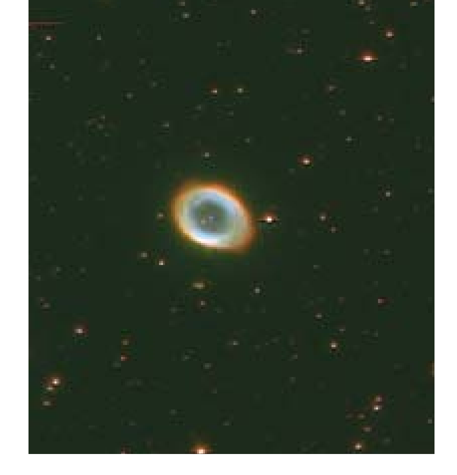
En su mayoría, las cámaras digitales se apoyan en los CCD para captar imágenes con igual procedimiento. Cuando la iluminación es baja, basta con que el CCD colecte luz durante un intervalo de tiempo mayor antes de cerrar el obturador. Desgraciadamente, esta táctica no funciona con un telescopio de espejo líquido, que no puede seguir a un objeto celeste en su movimiento por el firmamento. Si se usase el CCD de la manera habitual, las estrellas aparecerían constituidas en trazas, no en puntos nítidos.

Sin embargo, podemos utilizar de otro modo los dispositivos de acoplamiento de carga: manipulando los voltajes para que la imagen electrónica se desplace hacia el borde del detector *durante* la exposición. Alineando la dirección del desplazamiento de la carga con la dirección en que la luz estelar proyectada se mueve sobre el detector y aplicando voltajes a los electrodos del CCD al ritmo adecuado, se desplazarán los electrones a la misma velocidad que la imagen. Si este desplazamiento de carga se hace bien, no habrá emborronamiento de la imagen, pues



3. COMO TIENEN QUE GIRAR, los grandes espejos líquidos son propensos a las distorsiones. En el borde, donde el mercurio rota deprisa, el correspondiente movimiento relativo induce turbulencias en el aire circundante que afectan a la superficie. Esta imagen de prueba de un espejo líquido de 2,5 metros de diámetro, fabricado en la Universidad de Laval, se ha tomado con una técnica óptica que resalta las perturbaciones.









4. UNA DE LAS IMAGENES más impresionantes obtenidas con un telescopio de espejo líquido, muestra la Nebulosa del Anillo, en la constelación de Lira (*izquierda*). La tomó el Observatorio de Desechos Orbitales de la NASA, que captó también la de una pareja de galaxias (NGC 3424 y NGC 3430) en Leo Menor (*centro*); los segmentos rectos de luz son artefactos que las estrellas brillantes crean en el detector CCD. Igualmente impresionante es la imagen de una galaxia espiral sin bautizar (*derecha*) captada por el Gran Telescopio Cenital.

los electrones seguirán el paso de los fotones que los producen.

Cuando la imagen de la estrella alcanza el borde del CCD lo hacen también los electrones correspondientes, que se transfieren entonces al amplificador, que los mide. No se necesita obturador porque la lectura procede sin solución de continuidad, a una velocidad de decenas de líneas por segundo. La duración efectiva de la exposición es el tiempo que tarda la imagen estelar en cruzar la superficie del CCD: del orden de uno o dos minutos. Esta técnica de observación, el "barrido de deriva" (*drift scanning*), se emplea también en los telescopios corrientes; es una forma muy eficiente de obtener imágenes de una zona amplia de cielo.

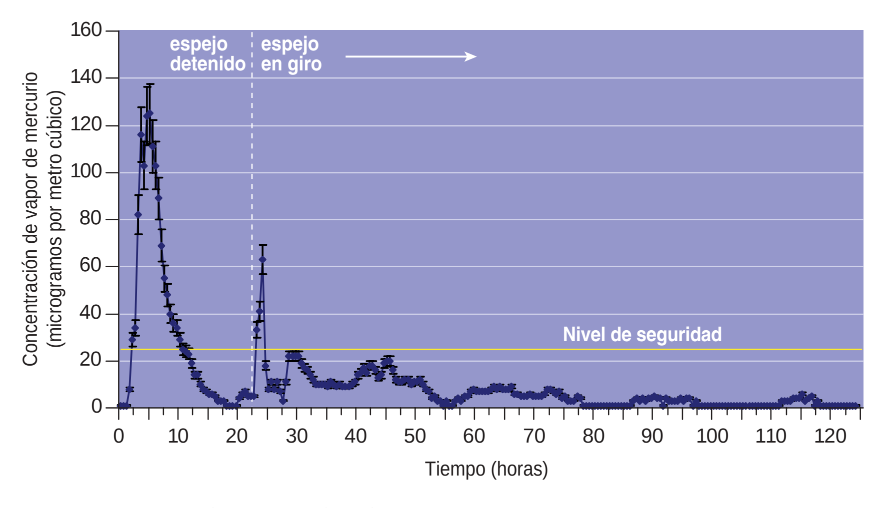
Aunque el barrido de deriva inspiró a Borra la recuperación de los telescopios de espejo líquido, no llegaría a aplicarlo. Sí se le deben otras innovaciones. Para evitar las vibraciones y fluctuaciones, recurrió a un cojinete de aire (una fina capa de aire presurizado entre dos superficies trabajadas con mucha precisión). Tales cojinetes, sin apenas rozamiento, proporcionan una rotación muy suave. Con un motor síncrono controlado por un oscilador cristalino, Borra eliminó las variaciones de velocidad que perjudicaron al instrumento de Wood.

En menos de diez años, Borra y sus colaboradores construían grandes espejos de mercurio, de hasta metro y medio de diámetro; más tarde harían uno de 3,7 metros. Esos espejos disponían de superficies de muy alta calidad óptica; el grupo logró en el laboratorio imágenes limitadas sólo por la difracción, esto es, imágenes tan nítidas como es teóricamente posible para un elemento óptico de ese tamaño. El físico canadiense utilizó incluso alguno de sus espejos en observaciones astronómicas, combinándolo con una cámara fotográfica de 35 milímetros.

Una gavilla de éxitos que atrajo a muchos otros astrónomos. El autor incluido. Tras duro trabajo, con mis colaboradores puse a punto en 1994 el primer telescopio de espejo líquido realmente práctico, en el que aplicamos el barrido de deriva. Aquel instrumento constaba de un espejo de mercurio de 2,7 metros de diámetro, una lente correctora de cuatro elementos para corregir la distorsión fuera de eje y una cámara CCD. Obtuvimos imágenes digitales de estrellas y galaxias con una resolución sólo limitada por el emborronamiento de la imagen causado por las turbulencias atmosféricas, no por imperfecciones del espejo líquido.

Al primero le siguieron pronto otros tres telescopios. Dos de ellos se usaron en investigaciones de la atmósfera de la Tierra por medio de láseres; con el tercero, el Observatorio de Desechos Orbitales, de 3 metros de diámetro, de la NASA, que operó en Nuevo México hasta 2002, se rastreó la "basura espacial".

Con la experiencia acumulada, mis colaboradores y yo fabricamos en



5. AUNQUE EL VAPOR DE MERCURIO es muy tóxico, los investigadores y el personal del observatorio corren poco peligro. El nivel de vapor tóxico sólo resulta preocupante cuando el espejo se detiene para drenar el mercurio durante la limpieza periódica, o cuando se vuelve a verterlo (*a la izquierda de la línea de puntos*). Pero en cuanto el espejo comienza a girar y la superficie del mercurio se recupera (*a la derecha de la línea de puntos*), se forma una fina capa de óxido que mantiene la concentración de vapores peligrosos por debajo del nivel que las normativas laborales consideran peligroso (*línea amarilla*).

CORTESIA DE MARK MULROONEY Y NASA (*Arriba a la izquierda y centro*), PAUL HICKSON (*arriba, derecha*); BARBARA AULICINO (*abajo*)/*American Scientist*

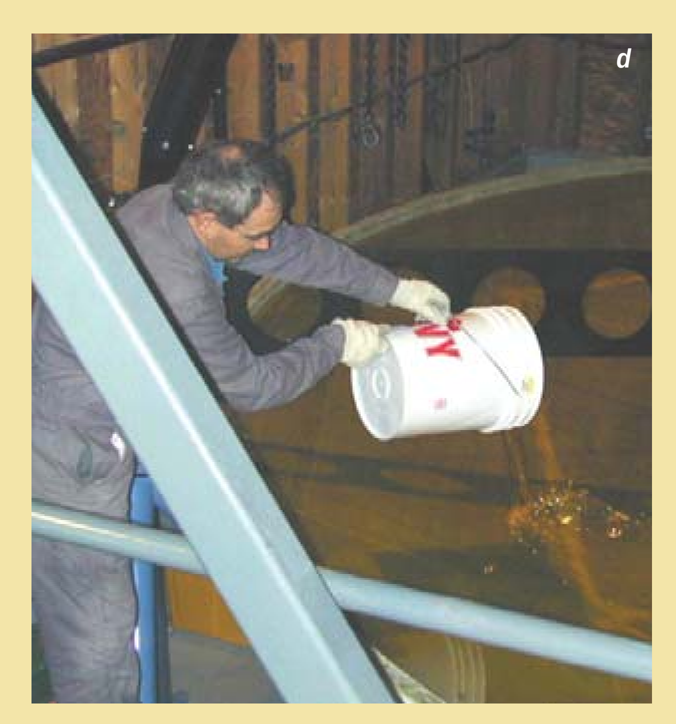




La maquinaria del Gran Telescopio Cenital

El ensamblaje del Gran Telescopio Cenital empezó en 1995 con la construcción del observatorio, levantado sobre lecho roqueño para limitar las vibraciones en el suelo. El suelo de hormigón tiene un metro de espesor. Las partes más bajas de las paredes, de hormigón también, están impermeabilizadas, como el suelo, de manera que puedan retener el peligroso mercurio en caso de fuga. En la base del telescopio se encuentra un cojinete, fabricado expresamente para el telescopio (*a*). Soporta el





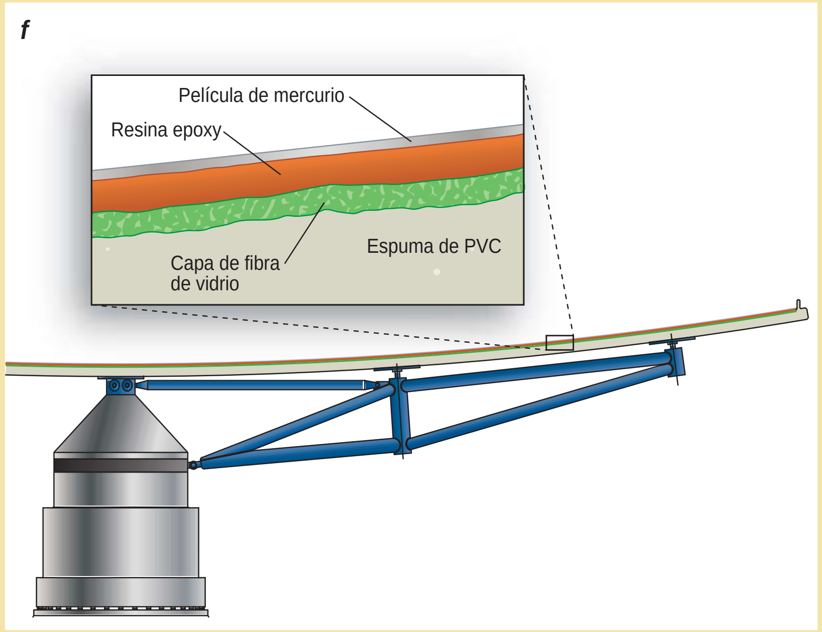
espejo de 3 toneladas sobre una delgada cámara de aire. Para eliminar cualquier perturbación, el compresor que aporta el aire del cojinete está a 20 metros de distancia, en otra habitación. Dentro del cojinete van un motor eléctrico sin fricción y un sensor óptico que detecta 2500 pulsos por revolución del espejo; así se controla la velocidad de giro con una precisión de una parte en un millón.

La parte superior del cojinete está fijada rígidamente a un entramado de tubos de acero (*b*), que soportan un gran disco de espuma ligera emparedada entre dos capas de plástico reforzado con fibra (*c*). El disco consta de siete hexágonos y seis triángulos de espuma, a los que se dio una suave curvatura calentándolos sobre una superficie, antes de recubrirlos y ensamblarlos. El entramado se fija al disco mediante monturas articuladas que pueden acomodarse a diferentes grados de expansión térmica de los materiales usados por encima y por debajo.

La espuma reforzada del disco adopta sólo de modo imperfecto la forma parabólica requerida. Se añadió una resina epoxy (*d*) para acercarse más a la forma deseada; esta resina, al principio líquida, se solidificó mientras el dispositivo rotaba. Sobre la resina se vertió una fina capa de mercurio, de unos 1,5 milímetros de espesor (*e*), que

PAUL HICKSON/*American Scientist*







proporciona al espejo una superficie reflectante de calidad óptica (*f*). Para recubrir el espejo de 6 metros de diámetro con esa capa se necesitan unos 40 litros de mercurio.

Una fina película de plástico mylar cubre la parte superior del espejo y rota con él. El mylar se fija a una estructura de soporte hecha de seis segmentos cuneiformes (*g*). Aunque esta tapa reduce un poco la cantidad de luz que incide sobre el espejo, elimina la formación de remolinos turbulentos de aire sobre el mercurio, que distorsionan la superficie reflectante.

MICHAEL DESJARDINS (*e*), BARBARA AULICINO (*f*); PAUL HICKSON (*g*)/*American Scientist*

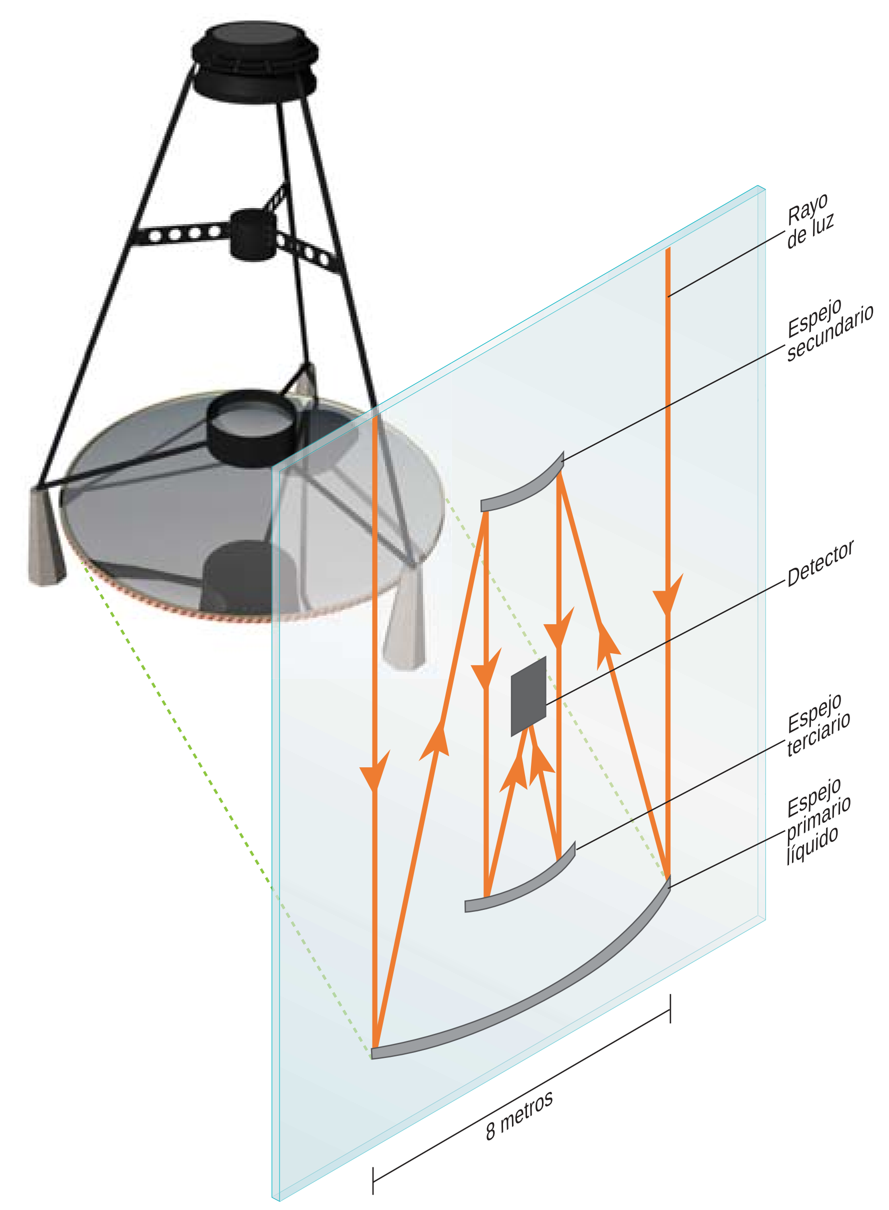
2005 un espejo de 6 metros para el Gran Telescopio Cenital. Perseguíamos desarrollar y perfeccionar espejos líquidos de un tamaño parecido al de los mayores telescopios. Por supuesto, la imposibilidad de observar con un telescopio de espejo líquido en una dirección diferente de la vertical (el cenit) significa que nunca desplazarán a los de espejo sólido, que apuntan a casi cualquier parte del cielo.

Pero esta falta de flexibilidad no limita tanto como podría imaginarse. En muchas investigaciones, apuntar a una misma dirección no es un impedimento. Si se quiere determinar las propiedades estadísticas de un número elevado de objetos distantes, de galaxias remotas, no importará, por lo general, qué parte del cielo se observe. Hay galaxias distantes por doquier y el cenit constituye un lugar tan idóneo como cualquier otro para buscarlas; más aún: es mejor, porque interpone la menor cantidad de aire en el camino óptico y, por tanto, causa la menor cantidad de absorción atmosférica, dispersión y distorsión de la imagen.

Para estudios de ese tipo, los telescopios de espejo líquido sirven igual que los corrientes y no son tan caros. El Gran Telescopio Cenital costó menos de un millón de dólares, un orden de magnitud por debajo de cualquier telescopio ordinario de tamaño equiparable. Si parte de los fondos destinados a la construcción de telescopios se gastase en instrumentos de espejo líquido, los astrónomos interesados en esas investigaciones dispondrían de bastante más tiempo de observación.

Baches en el camino

El éxito del Gran Telescopio Cenital tardó en llegar. Nos enfrentamos a innumerables obstáculos, la mayoría de ellos inesperados. Cuando aún estábamos empezando a diseñarlo, descubrimos que no había cojinete de aire que soportara nuestro espejo de seis metros y tres toneladas de peso. Hubo que encargar su fabricación. Y no resultó nada fácil. Lo logró una empresa de Minneapolis. Las dificultades con que tropezamos iban de lo predecible —fallos del equipo, problemas de sobrecalentamiento o congelamiento del sistema neumático que proporcionaba aire al



6. UN GRUPO DE INVESTIGADORES planea construir un gran telescopio de espejo líquido en la cima de una montaña chilena con óptimas condiciones para la observación. La ubicación sudamericana de esta "Sonda Avanzada de Espejo Líquido para Astrofísica, Cosmología y Asteroides" ha influido en la elección de su nombre (en inglés, el acrónimo es "ALPACA"). El telescopio tendría un espejo primario de ocho metros de diámetro, con espejos secundario y terciario de vidrio más pequeños y hasta 240 dispositivos de acoplamiento de carga en el detector.

BARBARA AULICINO/*American Scientist*

cojinete— a lo ridículo: un ratón saltó sobre el espejo giratorio y corrió en círculo sobre él, rompiendo la delicada capa de mercurio, pues le daba miedo saltar afuera.

De esos problemas, dos nos exigieron un considerable esfuerzo. El primero fue una reminiscencia de las fluctuaciones de velocidad que sufría el telescopio de Wood. Aunque nuestro espejo rota movido por un motor sincronizado a un oscilador muy estable, no hay un acoplamiento rígido entre el motor y el espejo, para evitar las consabidas vibraciones. Se conecta un anillo de imanes permanentes al rotor del cojinete de aire, la parte que soporta al espejo gigante. Estos imanes giran dentro del campo magnético de tres bobinas de campo estacionario. Proporcionando la corriente eléctrica adecuada a estas bobinas se genera un campo magnético rotativo, que a su vez hace que el espejo gire.

Creíamos que la poderosa inercia de nuestro espejo macizo impediría la aparición de corrientes de aire u otros fenómenos que produjeran variaciones en la velocidad de rotación. Andábamos errados. Al poner a girar el telescopio por primera vez, observamos que, en cuanto entraban golpes de viento por el tejado abierto del observatorio, la velocidad variaba hasta en un 0,1 por ciento: mil veces la inestabilidad de una parte en un millón que deseábamos.

Para salir del apuro proyectamos y construimos un sistema que estabilizaba la velocidad de rotación. Según gira el espejo, un dispositivo que lleva fijado genera por cada 360 grados de rotación 2500 pulsos espaciados con fina precisión. El sistema de control mide el momento en que llegan los pulsos y los compara con el momento en que deberían haber llegado si el espejo se hubiese movido a la velocidad prescrita. Si un pulso arriba tarde, indica que el espejo se está frenando: para compensarlo, debe incrementarse la potencia proporcionada a las bobinas estacionarias del motor. Si el pulso llega demasiado pronto, se ajustan las corrientes del motor para disminuir la velocidad del espejo. Así, se solucionó el problema. Ahora, la velocidad del espejo es constante en una parte por millón casi en cualquier circunstancia.

Descubrimos un segundo y más inquietante problema cuando comenzamos a tomar imágenes con el telescopio: las estrellas aparecían rodeadas por un halo de luz difusa. Ya se había visto este efecto en el laboratorio, pero nunca con tanta intensidad. Por la superficie del mercurio se propagaban numerosas ondas que difractaban la luz y emborronaban nuestras imágenes. Pero, ¿qué causaba las ondas?

Pronto lo descubrimos. El espejo líquido del Gran Telescopio Cenital es grande y rota bastante deprisa. El borde del espejo atraviesa el aire a una velocidad de 2,2 metros por segundo. Con respecto al mercurio de la parte externa del espejo, el aire se mueve a esa velocidad, razón por la que se producen olas, como cuando sopla el viento sobre el agua.

Sometimos a contrastación esta teoría midiendo el flujo de aire cerca de la superficie con espumillón sujeto en una varilla. Cerca del centro del espejo, el flujo de aire era suave, pero a un radio de un metro aproximado se volvía turbulento. Se formaban vórtices espirales a menos de un cen-

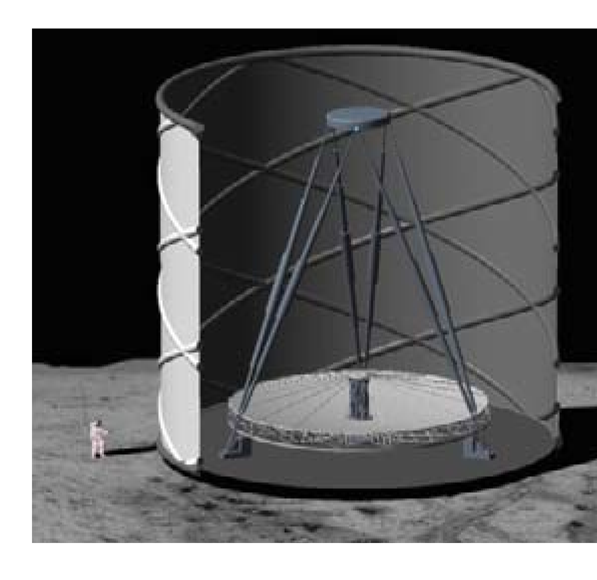
tímetro de altura sobre el mercurio, en el que creaban las ondas perturbadoras. El resultado se confirmó con pruebas ópticas minuciosas.

Nuestra solución fue suspender una capa muy fina de plástico mylar (una delgada lámina de poliéster) a escasos centímetros de altura sobre el mercurio. Esta cubierta rota con el espejo; atrapa el aire que tiene debajo y lo obliga a girar a la misma velocidad que el mercurio. Aún hay vórtices, pero ahora sobre el plástico; no pueden afectar a la superficie líquida reflectante. Una vez que añadimos esta cubierta, las ondas desaparecieron. El Gran Telescopio Cenital alcanza ahora una calidad de imagen que resiste la comparación con la alcanzada por los mejores telescopios ordinarios.

A la Luna

¿Qué deparará el futuro a los telescopios de espejos líquidos? Probablemente mucho. Ahora mismo están en marcha dos nuevos proyectos. El Telescopio Internacional de Espejo Líquido, a cargo de Jean Surdej, de la Universidad de Lieja, será un instrumento de 4 metros con una cámara CCD de gran formato. A diferencia del Gran Telescopio Cenital, que por necesidad se construyó cerca de mi facultad, en Vancouver, donde las condiciones de observación son mediocres, el nuevo telescopio se ubicará sobre una montaña chilena, un lugar soberbio para la astronomía. El segundo proyecto, dirigido por Arlin P. S. Crotts, de la Universidad de Columbia, se propone construir un telescopio de espejo líquido de 8 metros, también en Chile, la Sonda Avanzada de Espejo Líquido para Astrofísica, Cosmología y Asteroides (ALPACA). Tendría un campo de visión de 3 grados, con una cámara de barrido de deriva, dotado de 240 CCD. Si contasen con los fondos necesarios, empezaría a funcionar en apenas tres o cuatro años.

Las posibilidades futuras son aún más fascinantes. Con elementos correctores activos en el punto focal, un telescopio de espejo líquido podría seguir un objeto en el cielo durante 30 minutos ininterrumpidos, lo que permitiría operar en el infrarrojo. Es irrealizable en este momento porque las matrices de detectores infrarrojos no valen para el barrido de deriva. Los telescopios avanzados de espejo líquido podrían incluir óptica adaptativa para compensar la turbulencia atmosférica. Se podría también combinar la luz de varios espejos líquidos separados, según lo acostumbrado en los telescopios corrientes; así se obtienen sensibilidades y resoluciones muy elevadas.



7. CUANDO SE CONSTRUYAN telescopios de espejo líquido en la Luna podrían ofrecer el aspecto imaginado en esta ilustración. Un equipo dirigido por J. Roger P. Angel, de la Universidad de Arizona, está proyectando ya un telescopio lunar de espejo líquido.

CORTESIA DE THOMAS CONNORS, OBSERVATORIO STEWARD, UNIVERSIDAD DE ARIZONA/*American Scientist*

Quizá la idea más notable de las sugeridas hasta ahora sea la de construir telescopios de espejo líquido en la Luna. Nuestro satélite constituye un lugar ideal merced a su cielo oscuro, gravedad estable y ausencia de atmósfera que influya sobre el líquido y distorsione las imágenes. Situado cerca de los polos norte o sur de la Luna, un telescopio que apuntase al cenit observaría la misma zona del cielo durante meses; detectaría y estudiaría los objetos más remotos y débiles del universo visible. Sobre la Luna podrían desplegarse espejos líquidos rotantes ligeros de hasta 100 metros de diámetro. Estos reflectores gigantes no serían de mercurio, que se congelaría. Se emplearían líquidos iónicos de baja temperatura, que no se evaporan. Para volverlos reflectantes, se depositaría sobre su superficie una capa metálica, mientras el espejo girase. La Luna carece de atmósfera; un cojinete de aire no resultaría práctico: seguramente, el espejo flotaría sobre un campo magnético estabilizado por elementos superconductores.

El telescopio lunar de espejo líquido dista de ser una mera fantasía. Ya ha obtenido el respaldo de la NASA: un equipo internacional, encabezado por J. Roger P. Angel, de la Universidad de Arizona, está proyectando un instrumento así.

El autor

Paul Hickson, tras doctorarse en astrofísica por el Instituto de Tecnología de California en 1976, se trasladó a la Universidad de la Columbia Británica, donde ejerce como profesor en el departamento de física y astronomía. Ha participado en diversos comités asesores, entre ellos el de los telescopios gemelos Gemini y el del Telescopio Espacial James Webb. Preside el Comité Científico Asesor del Telescopio de Treinta Metros.

Bibliografía complementaria

WIDE-FIELD TRACKING WITH ZENITH-POINTING TELESCOPES. P. Hickson en *Monthly Notices of the Royal Astronomical Society*, vol. 330, págs. 540-546; 2002.

A LUNAR LIQUID-MIRROR TELESCOPE (LLMT) FOR DEEP-FIELD INFRARED OBSERVATIONS NEAR THE LUNAR POLE. S. P. Worden, J. R. P. Angel, D. J. Eisenstein, S. Sivanandam, J. H. Burge, E. F. Borra, C. M. Gosselin, O. Seddki, S. Thibault, P. Hickson, K. Ma, B. Fong y J. L. Josset en *Space Telescopes and Instrumentation I: Optical, Infrared, and Millimeter*, dirigido por J. C. Mather, H. A. MacEwen y M. W. M. de Graauw. SPIE—The International Society for Optical Engineering; Bellingham, Washington, 2006.

THE LARGE ZENITH TELESCOPE - A 6-METER LIQUID MIRROR TELESCOPE. P. Hickson, T. Pfrommer, R. Cabanac, A. Crotts, B. Johnson, V. de Lapparent, K. M. Lanzetta, S. Gromoll, M. K. Mulrooney, S. Sivanandam y B. Truax en *Publications of the Astronomical Society of the Pacific*. En prensa.

LOS USUARIOS DE UNA RED basada en codificaciones podrían tener muchas ventajas, entre ellas la de padecer menos demoras al descargar vídeos y al acceder a la Red.

Supresión de atascos en las redes de comunicaciones

Un nuevo método de transmisión de datos, la "codificación de red", podría elevar mucho la eficiencia de las redes de comunicaciones. Su idea central es sumamente curiosa: el envío de indicios sobre los mensajes puede resultar más útil que enviar los propios mensajes

Michelle Effros, Ralf Koetter y Muriel Médard





La historia de los modernos sistemas de comunicaciones ha estado marcada por fogonazos de pasmosa lucidez. Claude E. Shannon, matemático e ingeniero, puso en marcha una de tales revoluciones hace casi 60 años, al echar los cimientos de una teoría matemática de las comunicaciones, conocida hoy por "teoría de la información". Recogemos ahora los frutos de aquel trabajo —que se ocupaba de la compresión y la transmisión fiable de datos— en los actuales sistemas de telefonía, sean de tendido o inalámbricos, en los dispositivos de almacenamiento de datos —discos duros, CD, DVD y tarjetas de memoria flash— o en Internet.

Shannon abordó el problema de las comunicaciones por líneas telefónicas que transmitían a través de llamadas individuales. En nuestros días, la información viaja cada vez más por redes compartidas (caso de Internet), en las cuales una multitud de usuarios se comunican simultáneamente a través del mismo medio, sea éste un cable coaxial, una fibra óptica o, en el caso de los sistemas inalámbricos, el aire. Las redes compartidas ofrecen la posibilidad de mejorar la capacidad de servicio y el rendimiento de los sistemas de comunicaciones. También son causa de conflictos cuando existe competencia en el uso de recursos comunes. Pueden ser muchos los usuarios que demanden acceso a un servidor que ofrezca descargas de canciones o que deseen entrar en un "punto caliente" de conexión a un sistema inalámbrico.

El problema consiste, pues, en hallar procedimientos para que la compartición de recursos no provoque conflictos. A menudo, los operadores de redes tratan de resolver el problema aumentando los recursos, pero no pocas veces esa estrategia resulta insuficiente. Por ejemplo, aunque los tendidos de hilo de cobre, de cable coaxial o de fibra óptica pueden proporcionar una anchura de banda suficiente para uso doméstico o comercial, su instalación resulta onerosa y, una vez instalados, cuesta modificarlos o ampliarlos. Los sistemas de transmisión multiantena de banda ultraancha pueden aumentar el número de usuarios que reciben servicio de las redes inalámbricas; aun así, quizá resulten insuficientes para satisfacer una demanda que crece sin cesar.

Son necesarias, pues, técnicas que aumenten la eficiencia de los sistemas. En Internet y en otras redes compartidas, la información pasa por una serie de "enrutadores" que actúan a modo de semáforos o cambios de agujas en los nodos donde se cruzan los caminos, los enlaces (*links*), que puede tomar la señal. Los enrutadores dirigen los mensajes hacia enlaces que conducen hasta sus destinatarios finales. Pero si lo que se busca es mejorar el rendimiento, ¿serán los enrutadores los dispositivos óptimos para resolver el paso por los cruces? Y más todavía, ¿será la conmutación el procedimiento idóneo?

Siete años atrás, esas cuestiones no se planteaban. Hasta que, en 2000, Rudolf Ahlswede, Ning Cai, Shuo-Yen Robert Li y Raymond W. Yeung dieron a conocer un nuevo método para la distribución de información a través de redes compartidas. En su propuesta, llamada "codificación de red", los enrutadores son sustituidos por codificadores que, en lugar de los mensajes propiamente dichos, envían indicios de cuál puede ser su contenido. Los equipos receptores deducen el mensaje original ensamblando los indicios que reciben.

Aunque este proceder pueda antojarse contrario a la intuición, la codificación de red, que todavía está en estudio, ofrece la esperanza de incrementar de forma impresionante la velocidad y la fiabilidad de todo tipo de sistemas de comunicación. Parece posible que sea la chispa que encienda la próxima revolución en este campo. Se están investigando otras vías para aumentar el rendimiento, aunque se trata, por lo general, de ampliaciones de métodos ya existentes.

Los bits no son coches

El grupo encabezado por Ahlswede elaboró su propuesta basándose, en parte, en una idea que se le debe a Shannon, a saber: que la transmisión de indicios de los datos puede resultar más útil que la transmisión de los datos en sí. Un receptor puede deducir los datos originales en cuanto recopile suficientes indicios sobre ellos; no tiene por qué recibir la totalidad de los indicios emitidos. Se puede reemplazar un tipo de indicio por otros; lo que importa es que se reciba una cierta combinación de indicios que revele el mensaje original. (Los receptores habrán de estar informados previamente sobre las reglas utilizadas para generar los indicios; o bien los propios indicios habrán de contener instrucciones sobre la forma de utilizarlos.)

La codificación de red rompe con la concepción clásica de unos canales de comunicación asimilados a carreteras y unos bits asociados a automóviles circulantes. Conviene,

no obstante, comprender el símil de la red de carreteras para entender el funcionamiento de la codificación de red y por qué es tan prometedora.

Shannon recurrió a la matemática para demostrar que cada canal tiene una capacidad —el caudal de información que puede circular por él durante un intervalo de tiempo determinado— y que las comunicaciones proseguirán de manera fiable en tanto no se desbordare. En la analogía con el transporte, la capacidad de una vía es el número máximo de vehículos por segundo que transitan sin tropiezos por ella. Mientras el tráfico se mantenga por debajo de la capacidad de la vía, a los vehículos que entren en ella se les podrá garantizar que saldrán por el otro extremo intactos (salvo accidentes de naturaleza excepcional). Este modelo ha inspirado sistemas de comunicaciones cada vez más complejos. Así, los sistemas telefónicos que estudiaba Shannon dedican una “vía” diferente para cada conversación; en las líneas telefónicas tradicionales nunca ocurre que dos llamadas distintas compartan una misma línea al mismo tiempo y en la misma frecuencia.

Las redes de ordenadores, e Internet en concreto, constituyen, en esencia, un laberinto de vías que se ramifican, se cruzan y confluyen. La información que viaja de un ordenador a otro lo hace, normalmente, por varias “carreteras”. Los bits de un mismo mensaje se agrupan en paquetes (los autobuses de las autopistas de la información), cada uno de los cuales está rotulado con la dirección del receptor al que va destinado. Los enrutadores, que se encuentran en las encrucijadas, examinan el rótulo de cada paquete de datos y le indican la ruta hacia su destino.

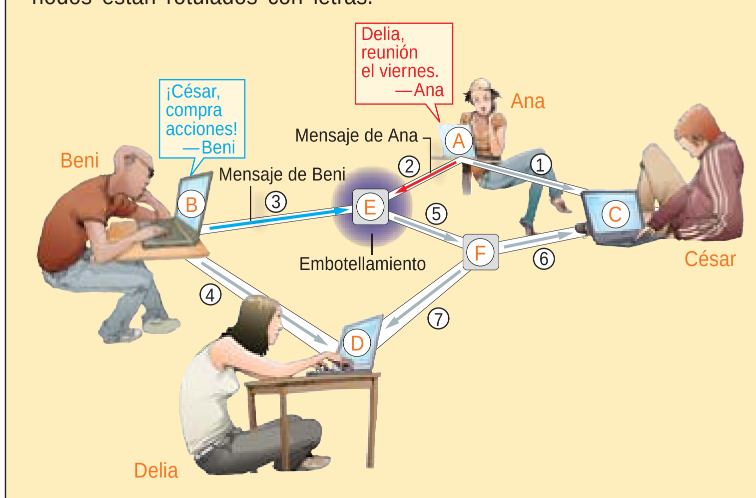
No deja de resultar paradójico que el modelo de transporte en el que se han inspirado los complejos sistemas de comunicaciones de nuestros días esté obstaculizando la senda del progreso. Los bits no son coches. Cuando dos vehículos se encuentran ante un mismo paso estrecho, han de atravesarlo por turno. Pero cuando dos bits llegan a un angostamiento, caben otras opciones. Ahí interviene la codificación de red.

Resumen/*Codificación de red*

- En el año 2000 se propuso una idea aparentemente absurda para reducir los atascos en las redes de comunicaciones, pero que quizá resulte revolucionaria: la “codificación de red”. Reemplaza los enrutadores, que en los nodos se limitan a retransmitir mensajes, por codificadores, que en lugar de retransmitir los mensajes propiamente dichos envían indicios relativos a los mensajes que les llegan.
- La codificación de red está desenvolviéndose con éxito en los experimentos, la mayoría de los cuales se centran, de momento, en el envío de datos por redes de multidifusión, donde todos los usuarios reciben simultáneamente la misma información.
- Este sistema promete mejorar el rendimiento de muchas redes, aumentar su capacidad sin tener que añadir equipos ni ampliar su ancho de banda y hacer que sean más rápidas, fiables y resistentes a los ataques.

LA IDEA FUNDAMENTAL

Una red sencilla, de sólo seis nodos, en la que los mensajes viajan por enlaces (*numerados*) a razón de 1 bit por segundo, permitirá ilustrar en qué consiste la codificación de red. Los nodos están rotulados con letras.



EN UNA RED ORDINARIA

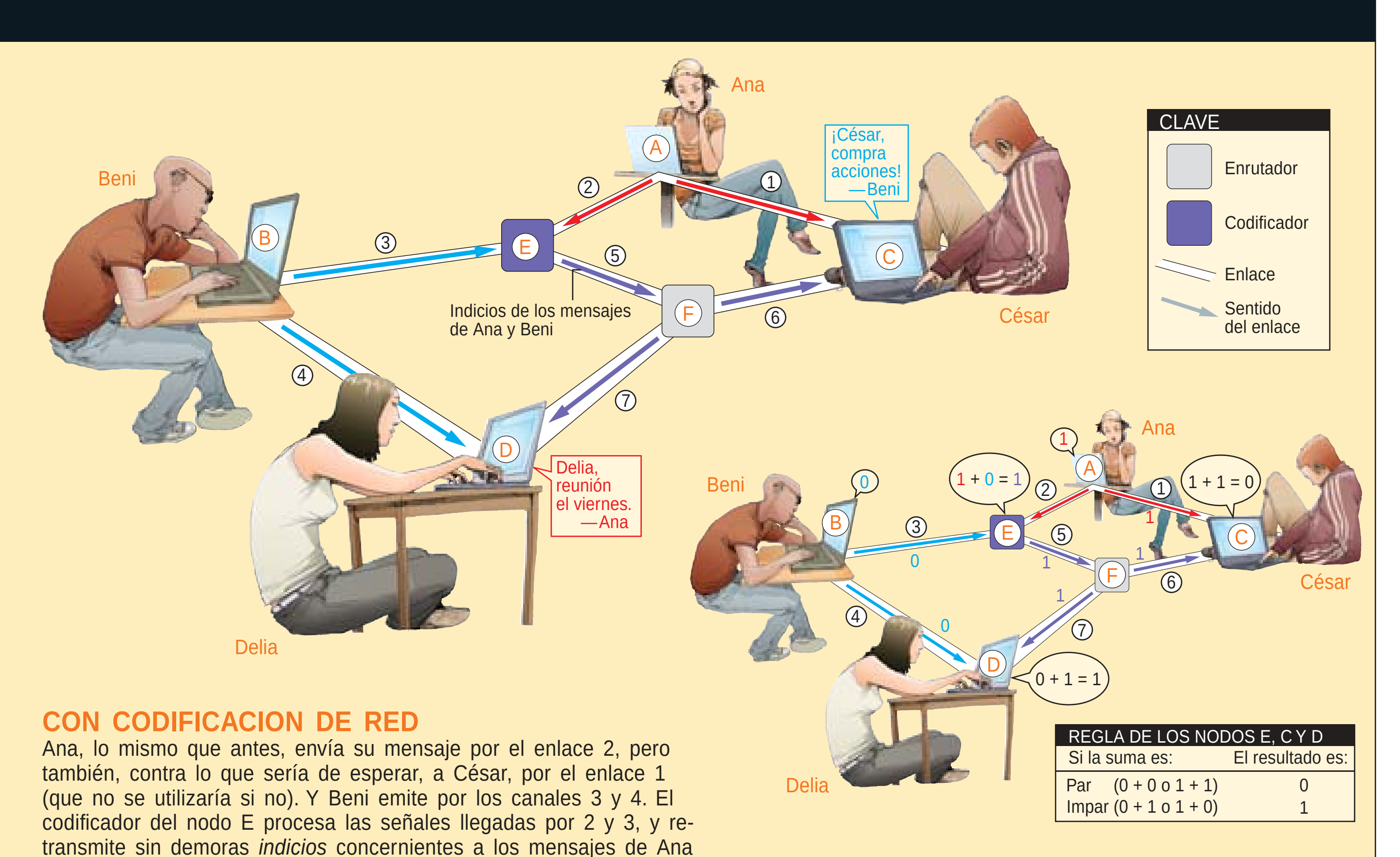
Ana, en el nodo A, le envía un mensaje a Delia, en el nodo D, justamente cuando Beni, desde el nodo B, manda uno a César, en C: los mensajes se atascarán en el nodo E, donde un enrutador tendría que ir retransmitiendo los mensajes, a razón de uno por vez, a través del enlace 5. Hay demoras.

MATT VINCENT

¿Cómo funciona?

La hipotética red digital de seis nodos que se ha representado en el recuadro “La idea fundamental” nos ayudará a aclarar esas opciones. Recordemos que, en los ordenadores, todos los mensajes adoptan la forma de “cadenas,” largas ristras de código binario. Imaginemos que cada enlace —cada vía— de esta red pueda transportar 1 bit por segundo (ya sea un 0 o un 1) y, además, solamente en la dirección indicada por la flecha correspondiente. Ana, una de las usuarias, está situada en el nodo A y espera poder enviar información a Delia, que está en el nodo D, a razón de 1 bit por segundo. Entretanto, Beni, desde el nodo B, confía en enviar, exactamente en el mismo momento y con la misma velocidad, información a César, en el nodo C. ¿Será posible satisfacer simultáneamente las demandas de Ana y de Beni sin exceder las capacidades del enlace?

En un sistema ordinario, que se valdrá de enrutadores (*véase el recuadro* “La idea fundamental”, *izquierda*), las perspectivas son desalentadoras. Los dos itinerarios (de Ana a Delia y de Beni a César) pasan por la vía 5. Este enlace viene a ser como un puente estrecho y de un solo carril. El enrutador del nodo E, donde comienza el enlace 5, recibe un total de 2 bits por segundo (uno procedente del enlace 2, el otro del enlace 3); ahora bien, como la capacidad del enlace 5 es de 1 bit por segundo, el enrutador sólo puede enviar por él un bit cada segundo. En el modelo de transporte, estos estrechamientos —”cuellos de botella”— provocan atascos de auténtica pesadilla, al irse apilando en ellos cada vez más bits con el paso del tiempo, en cola hasta que les llegue el turno.



REGLA DE LOS NODOS E, C Y D	
Si la suma es:	El resultado es:
Par (0 + 0 o 1 + 1)	0
Impar (0 + 1 o 1 + 0)	1

CON CODIFICACION DE RED

Ana, lo mismo que antes, envía su mensaje por el enlace 2, pero también, contra lo que sería de esperar, a César, por el enlace 1 (que no se utilizaría si no). Y Beni emite por los canales 3 y 4. El codificador del nodo E procesa las señales llegadas por 2 y 3, y retransmite sin demoras *indicios* concernientes a los mensajes de Ana y de Beni por la línea 5. Esos indicios se reenvían a continuación por los enlaces 6 y 7. El ordenador de César descifra el mensaje de Beni examinando el mensaje de Ana más los indicios recibidos por el enlace 6, y el ordenador de Delia descifra el mensaje de Ana procediendo de forma análoga. ¿Cómo es posible? Siga leyendo.

LA DESCODIFICACION

Imagine que el mensaje de Ana es el número 1, que el de Beni es un 0 y que el codificador emplazado en E, así como los ordenadores de César y Delia, procesan dos flujos de datos de entrada conforme a la regla de codificación expuesta en la tabla. El nodo E enviaría un 1 por el enlace 5 y el ordenador de César combinaría el 1 recibido por el enlace 1 con el 1 recibido por el enlace 6; se obtendría un 0 y se recuperaría el mensaje de Beni. Mientras, el ordenador de Delia combinaría el 1 del enlace 7 y el 0 del enlace 4 y produciría un 1, que es el mensaje de Ana. En la realidad, si los nodos C y D no estuvieran previamente programados con la regla, el codificador del nodo E la transmitiría hasta ellos junto con el indicio que envió por el enlace 5.

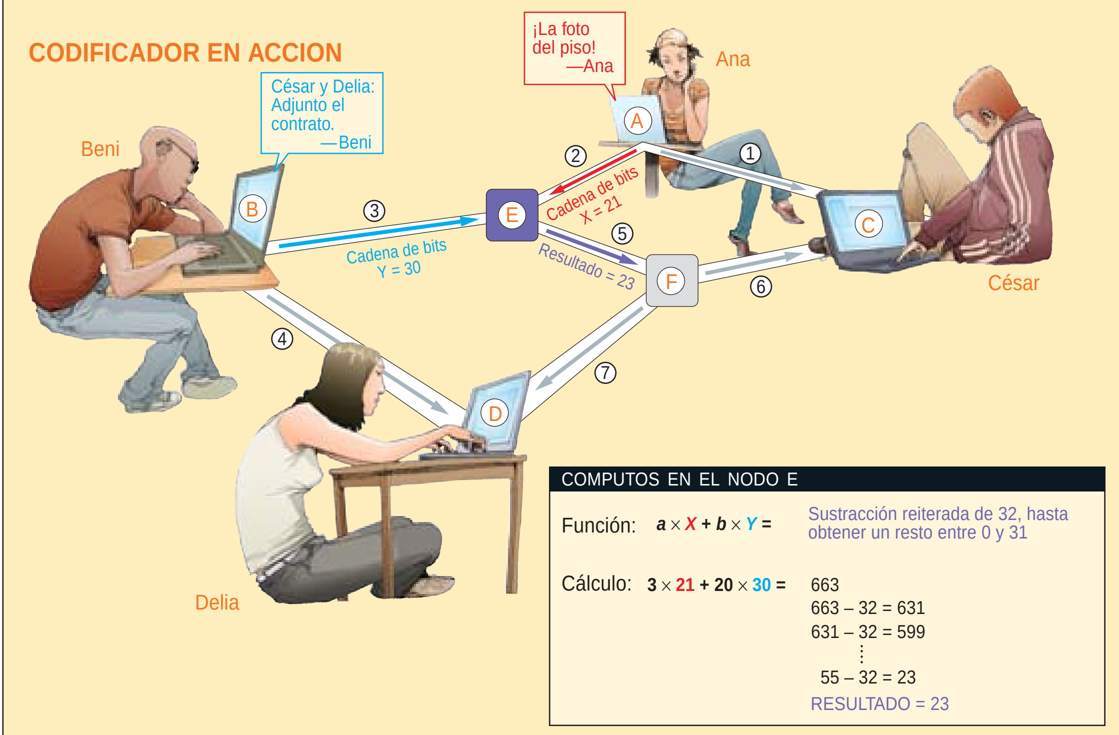
En el nuevo procedimiento (*véase el mismo recuadro a la derecha*), el enrutador típico ser reemplazaría por un codificador, que dispondría de más opciones que un mero "agente de tráfico". En vez de ir retransmitiendo sin más los flujos de datos que, esperando su turno, se han apilado en el angostamiento, el codificador podría enviar una información muy diferente. Podría, por ejemplo, contar el número de unos que le llegan durante un segundo determinado y transmitir un 0 si es par. Si el número fuese impar, el dispositivo transmitiría un 1. De ese modo, si el enlace 5 recibe simultáneamente un 1 y un 0, llegados por los enlaces 2 y 3, lo que hace es cursar un 1. En cambio, si por los enlaces 2 y 3 le llegan simultáneamente ora dos 0, ora dos 1, el enlace 5 transportará un 0. A continuación, este resultado se enviará por el enrutador F a través de los enlaces 6 y 7 hasta César y Delia, respectivamente.

Con este método, cada par de bits del nodo E es reemplazado por un híbrido de los dos. Un flujo de bits tal parecerá absurdo. El codificador que proponemos realiza el equivalente a cruzar entre sí dos conversaciones telefónicas, volviéndolas más difíciles de entender. La apariencia absurda de tal proceder ha sido la causa de que permaneciera sin investigar durante tanto tiempo.

En ocasiones, lo que parece locura resulta ser innovación. Un flujo híbrido de bits no puede describir a la perfección ninguna de las dos transmisiones, aunque sí proporcionar indicios acerca de ambas. Supongamos que además le fueran enviadas, a César, la misiva de Ana por el enlace 1, y a Delia, por el enlace 4, el mensaje de Beni. Para enviar estos dos mensajes se utilizan recursos de la red (enlaces 1 y 4) que el sistema enrutador no podía utilizar para atender las peticiones de Ana y de Beni. El nodo de César recibe la transmisión enviada por Ana y sabe en cada instante (gracias al enlace 6) si el número de unos en el par de mensajes emitidos por Ana y Beni es par o non. Si el nodo de César se programa para que también "conozca" la regla utilizada por los codificadores en la entrada del enlace 5, o si es capaz de inferir la regla a partir de los propios indicios,

UN EJEMPLO DE DISEÑO

Para hacernos más clara idea de cómo procesan información los codificadores y para ver por qué los proyectistas de redes de multidifusión pueden seleccionar fácilmente las reglas que van a utilizar los codificadores, volvamos a examinar la red de seis nodos. Esta vez, sin embargo, Ana y Beni desean enviar información simultáneamente a César y a Delia. Se trata de un problema de multidifusión, porque ambos receptores esperan poder reunir todos los datos.



CADENAS BINARIAS DE 5 BITS

Cadena m = 5	Valor numérico	Valor en el ejemplo
00000	0	
00001	1	
00010	2	
00011	3	*a*
00100	4	
00101	5	
00110	6	
00111	7	
01000	8	
01001	9	
01010	10	
01011	11	
01100	12	
01101	13	
01110	14	
01111	15	
10000	16	
10001	17	
10010	18	
10011	19	
10100	20	*b*
10101	21	*X*
10110	22	
10111	23	
11000	24	
11001	25	
11010	26	
11011	27	
11100	28	
11101	29	
11110	30	*Y*
11111	31	

En este caso, el problema de la codificación se reduce, en esencia, a la selección de una regla de codificación —una función matemática— para el nodo E, que ingresa información por los enlaces 2 y 3 y transmite el resultado de la función codificadora por el enlace 5 (*arriba*). Hay muchas funciones posibles aptas para nuestro propósito, pero nos fijaremos en las funciones lineales para hacer ver lo muy sencillos que pueden ser los códigos. Además, como la codificación de red se aplica normalmente a bloques de bits, diseñaremos la función correspondiente al enlace 5 como una combinación aleatoria de cierto conjunto numérico (m) de bits procedentes del enlace 2 y de un número igual de bits tomados del enlace 3.

Uno de los métodos posibles consiste en tomar $m = 5$. Existen 32 cadenas binarias de longitud 5, que representaremos mediante un entero de 0 a 31 (*tabla de la derecha*). En el nodo E podemos combinar las cadenas mediante las funciones $a \times X + b \times Y$, en donde X e Y son los valores entre 0 y 31 que corresponden a las ristras que el nodo E recibe de los enlaces 2 y 3 (sean, por ejemplo, 21 y 30 en un determinado instante) y donde a y b son también ristras binarias de longitud m (arbitrariamente hemos elegido que sean 3 y 20). Nos aseguramos de que el resultado de esta combinación lineal sea también un elemento del conjunto finito de 0 a 31. Ello se consigue efectuando las operaciones de la fórmula y restando 32 del resultado las veces que sea necesario hasta lograr un entero comprendido entre 0 y 31. Este último número es el resultado que arroja el codificador.

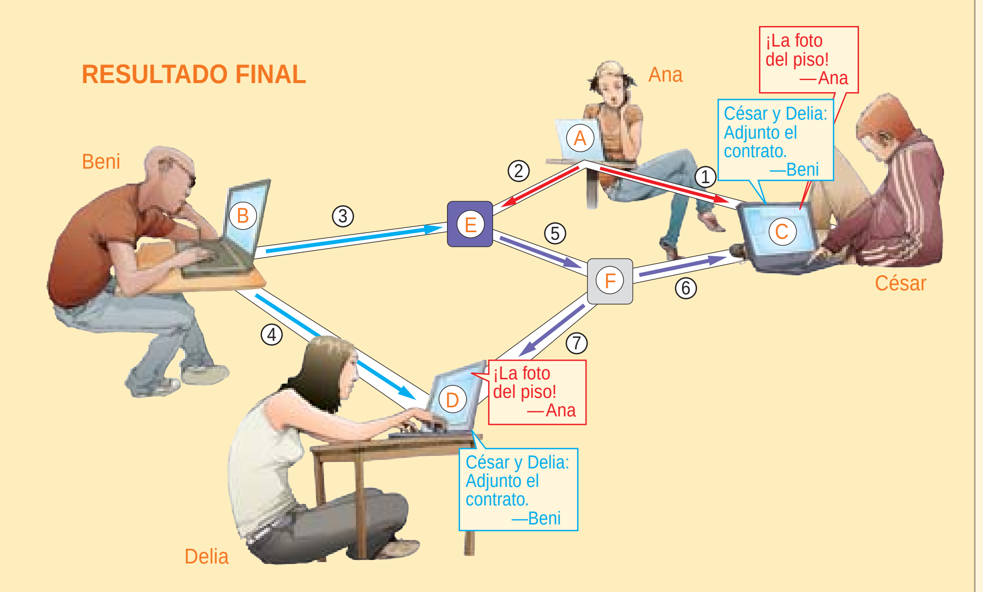
los que haya recibido lo capacitarán para descifrar el mensaje enviado por Beni. Y de forma similar, el nodo de Delia podrá desvelar el mensaje de Ana.

Ventajas claras

Esta estrategia logra dos metas que eran impensables, dadas las limitaciones del modelo del transporte. En primer lugar, permite que el bit que parte de un nodo recorra dos caminos simultáneamente, algo que no puede hacer un coche. En segundo lugar, permite que un par de flujos de bits que llegan a la entrada de un cuello de botella se combinen en un solo flujo, mientras que dos coches que concurran en un puente estrecho no pueden convertirse en una entidad: uno de ellos tendría que esperar a que pasara el otro antes de cruzar el puente.

El método de manejo de datos expuesto en nuestro modelo con seis nodos (una pequeña modificación del que Ahlswede y sus colegas presentaron en 2000) ofrece la posibilidad de aumentar la capacidad de una red sin exigir la adición de conducciones nuevas, porque evita los atascos. Si sólo utilizara enrutadores, nuestra red de seis nodos podría sostener transmisiones simultáneas que, en promedio, equivaldrían a medio bit por segundo. (Debido a que las dos transmisiones en competencia tendrían que compartir el enlace 5, la tasa de bits efectiva sería de un bit cada dos segundos para cada una de las demandas en conflicto, es decir, medio bit por segundo.) Con la codificación de red, el mismo sistema soporta transmisiones simultáneas al ritmo de un bit por segundo. Así pues, en este ejemplo, la codificación de red duplica la capacidad.

MATT VINCENT



Si Ana le envía a César su mensaje por el enlace 1 (además de enviarlo por el enlace 2 hasta el nodo E) y si Beni envía su mensaje por el enlace 4 así como por el enlace 3 (*arriba*), los ordenadores de César y Delia podrán insertar la salida de *E* en la misma función aplicada por *E* y deducir *X* e *Y*, respectivamente. La repetición del mismo proceso para cada ristra de bits acaba revelando ambos mensajes.

FACIL ELECCION DE CODIGOS

A primera vista, la selección al azar de códigos podría no parecer buena idea. Dado que a y b pueden tomar cada uno cualquier valor en el conjunto de 0 a 31, son $32 \times 32 = 1024$ las posibles funciones de la forma $a \times X + b \times Y$ que podríamos utilizar en el nodo E, y no todas esas funciones van a asegurar que ambos receptores puedan determinar los mensajes originales X e Y. Por ejemplo, si utilizamos un código en el que $a = 0$, Delia recibe Y en el enlace 4 y $b \times Y$ en el enlace 7 y esta información no será suficiente para obtener el mensaje X, independientemente de cuál sea el valor de b.

Sin embargo, cuando ninguno de los coeficientes a o b es 0, ambos receptores pueden recuperar con éxito los mensajes debidos. En consecuencia, de las 1024 funciones posibles todas menos 63 son apropiadas para la codificación en el nodo E. Aun cuando los coeficientes se elijan al azar, se obtiene un código aceptable en un 94 por ciento de las veces. La probabilidad de éxito se aproxima rápidamente al 100 por 100 al aumentar la longitud m de las cadenas, por lo que todo cuanto han de hacer los programadores para asegurar la fiabilidad es elegir un valor de m adecuadamente grande.

En algunas ocasiones, la codificación de red proporcionaría ganancias en capacidad todavía mayores, en otras, ninguna. Pero el sistema nunca reduciría la capacidad de la red, porque, en el caso más desfavorable, lo que haría sería remedar las acciones de los sistemas con enrutadores. La codificación de red debería también aumentar la fiabilidad y la resistencia a los ataques en redes esenciales, pues, gracias a la naturaleza intercambiable de los indicios, la pérdida de algunos paquetes de ese tipo de información no causaría problemas.

Lecciones de las redes multidifusión

Hasta ahora, la investigación sobre la forma de implantar redes codificadas se ha centrado, sobre todo, en las redes de multidifusión; en éstas, todos los receptores han de recibir la misma información. Los videojuegos de Internet se basan en sistemas de multidifusión, pues han de poner al corriente a todos los jugadores cada vez que uno de ellos efectúa un movimiento. Las retransmisiones por Internet de vídeos o acontecimientos deportivos, o la distribución electrónica de nuevos programas informáticos a un número elevado de clientes, viajan también por redes multidifusión. En la actualidad, tales redes siguen utilizando enrutadores. Nos resultará más fácil explicar por qué su diseño suele resultar tan difícil volviendo al símil de transporte por carretera.

Imaginemos que todas las carreteras del país estuvieran llenas de autobuses (los paquetes de datos). Los enrutadores vendrían a ser agentes de tráfico que dirigirían la circulación en cada cruce. Los autobuses que fueran llegando se sumarían a la cola, detrás de los vehículos que les precedieron en la arribada. El agente iría leyendo en cada autocar a qué destino se encamina, y le dirigiría hacia la vía que ha de seguir.

En el diseño de sistemas, se busca lograr que cada enrutador dirija el tráfico de forma que no sólo haga viajar más rápidamente a cada coche hacia su destino previsto, sino también que el sistema nacional de carreteras, en su conjunto, atienda a tantos conductores como sea posible.

Incluso a un proyectista que dispusiera de una autoridad central y de un mapa completo de las carreteras del país, le resultaría difícil determinar la estrategia óptima que cada enrutador habría de seguir. La dificultad aumenta al ir variando con el tiempo el estado de la red: horas punta, reparaciones de las carreteras, accidentes o acontecimientos deportivos llevan consigo que tanto las carreteras como las prestaciones que se les exigen se hallen en permanente cambio.

La intuición podría hacer pensar que el diseño de una red basada en la codificación resultaría más difícil, porque habría más opciones que considerar. Un cierto nodo podría estar dejando pasar los datos sin modificación, remedando así el funcionamiento de un enrutador. Pero también podría encontrarse mezclando dos o más flujos de datos antes de darles salida. La propia forma en que los mezcla podría ser objeto de consideración. Más todavía: nodos diferentes podrían utilizar distintos algoritmos de codificación.

Breve historia de la codificación de red

Las entradas siguientes recuerdan algunos hitos; la bibliografía complementaria ofrece otras referencias. Hay una bibliografía más completa en **www.ifp.uiuc.edu/~koetter/NWC/**

2000: Presentación de la idea. Rudolf Ahlswede, Ning Cai, Shuo-Yen Robert Li y Raymond W. Yeung hicieron ver la potencia de la codificación de redes en el caso de la multidifusión, que envía información idéntica a todos los receptores. Esos autores demostraron la existencia de códigos idóneos (informativos), pero no expusieron un método para diseñarlos.

2003: Avances importantes hacia la realización práctica. Li, Yeung y Cai demostraron que la codificación de redes de multidifusión puede basarse en funciones matemáticas que sólo utilicen adición y multiplicación, lo que reduce la complejidad del diseño de códigos. Dos de los autores (Koetter y Médard) introdujeron un armazón algebraico para el análisis de métodos de codificación y la simplificación del diseño de códigos.

2005-2006: Publicación de valiosos algoritmos de diseño. Sidhart Jaggi, del Instituto de Tecnología de California (Caltech), con Peter Sanders, de la Universidad de Karlsruhe, uno de los autores (Effros) y colaboradores, y por otra parte Tracey Ho, del Caltech, con los tres autores y otros colaboradores, publicaron algoritmos de poca complejidad para diseñar las funciones utilizadas en cada nodo de una red de multidifusión. El primero de estos artículos proporcionaba una metodología sistemática para el diseño de funciones; el segundo hacía ver que la elección independiente y aleatoria de funciones en cada nodo debería funcionar con similar eficacia. (Las primeras versiones de ambas contribuciones se presentaron en un congreso de 2003.)

2006: Exploración de aplicaciones para redes inalámbricas. En un congreso celebrado en 2006, el grupo encabezado por Christina Fragouli, de la Escuela Politécnica Federal de Lausana, expusieron las ventajas potenciales de la codificación para aplicaciones en redes inalámbricas y describieron supuestos en los que ese enfoque resultaría de especial utilidad.

MATT VINCENT

Felizmente, estos razonamientos son erróneos. En ocasiones, al añadir más opciones se simplifican las cosas. Sin codificación, los arquitectos de un sistema de multidifusión tendrían que enumerar tantos itinerarios desde el emisor hasta cada uno de los receptores como fuera posible, para determinar después cuántas de estas rutas podría soportar la red simultáneamente. Incluso en el caso de redes sencillas, la determinación y ensayo de todas las combinaciones de caminos constituiría una tarea enloquecedora.

En contraste, el diseño de un sistema de multidifusión que utilizase la codificación de red resultaría bastante sencillo. La asombrosa verdad es que las únicas operaciones matemáticas que se necesitan para la codificación de redes son la adición y la multiplicación. Y todavía más: si la función o regla programada en cada codificador de una red es elegida con independencia del mensaje y de las otras funciones de codificación, y sin conocimiento alguno de la organización de la red, la probabilidad de que el sistema completo funcione con un rendimiento máximo es muy elevada. Aunque el sistema vaya cambiando con el tiempo, pensemos en lo que ocurre en las redes móviles o configurables, la red seguirá funcionando de forma óptima sin necesidad de rediseñarla. Para comprender por qué, véase el recuadro “Un ejemplo de diseño”.

Las redes de mañana

El funcionamiento de las redes sería, pues, muy diferente si los codificadores sustituyeran a los enrutadores. Los mensajes viajarían por las redes de otro modo: no sólo compartirían la “autopista” con otras transmisiones, sino que se imbricarían con tráfico procedente de múltiples fuentes. Hay quienes temen que esta “intimidad” pueda comprometer la seguridad de los mensajes. Lo más probable es, sin embargo, que el tráfico que circule por las redes se convierta en un flujo algebraico localmente indescifrable. Los usuarios de la red colaborarían sin saberlo, para mutua ventaja de unos y otros, y así no sólo se lograrían tasas más elevadas de transferencia o descarga de datos, sino también, en el caso de las redes inalámbricas, una mejora de la eficiencia energética. (Dado que cada emisión inalámbrica consume energía, se podría reducir el consumo de los nodos mezclando entre sí la información destinada a varios vecinos y enviando después una sola emisión.)

Además, las demoras en las descargas de vídeo y las llamadas perdidas de los teléfonos móviles serían mucho menos frecuentes. En Internet, los enrutadores fallan, o quedan fuera de servicio por tareas de mantenimiento, y se pierden sin cesar paquetes de datos. Por eso es necesario recargar páginas de la Red y hay páginas que bajan muy lentamente. La fiabilidad aumentaría con la codificación de red, porque no son imprescindibles todos y cada uno de los indicios para que funcione el sistema.

Y los gestores de las redes podrían proporcionar todas estas ventajas sin tener que añadir nuevos canales de comunicaciones, porque se haría mejor uso de los canales ya existentes. La codificación de red sería de este modo un complemento de otras técnicas de

comunicaciones y permitiría que los usuarios obtuvieran el máximo partido de ellas.

En ocasiones, los usuarios sabrían que la codificación de red está actuando porque modificaría el funcionamiento de algunas aplicaciones de uso frecuente, como las descargas *peer to peer* (P2P, transferencias entre iguales). En la actualidad, cuando alguien desea descargar un archivo por este sistema lo que se hace es indagar si existe algún usuario en cuya máquina resida el archivo y esté dispuesto a colaborar. En un sistema que utilizase la codificación de red, el archivo ya no se almacenaría en su conjunto ni en piezas reconocibles.

Pero los usuarios no tendrían que averiguar personalmente cómo hallar los indicios necesarios para obtener los archivos deseados. Una petición enviada a una red desde un teléfono o desde el ordenador de un usuario provocaría que, bien el ordenador del solicitante, bien un servidor local, rebuscase por la red indicios relacionados con el archivo de interés. La información que así se reuniese, consistente en piezas algebraicamente mezcladas de información relativa al archivo deseado, ayudaría a recuperarlo. En lugar de ensamblar un rompecabezas cuyas piezas serían fragmentos reconocibles de un todo, el servidor, o el ordenador del solicitante, resolvería un sistema de ecuaciones algebraicas. Y durante el proceso, casi todos los participantes permanecerían en bendita ignorancia de estas operaciones, tal y como casi todos desconocemos las complicadas operaciones de corrección de errores de nuestros teléfonos móviles.

El ejército ha reconocido la robustez de la codificación de red. El de Estados Unidos financia la investigación para utilizarla en redes locales *ad hoc*, que se crean sobre la marcha. Tales redes adquieren extraordinario valor en ambientes sujetos a cambios rápidos, como el campo de batalla, donde el mantenimiento de comunicaciones fiables es esencial y donde el montaje y mantenimiento de una estructura de fibra óptica, cables coaxiales o torres de antenas resulta difícil. En una red constituida "sobre la marcha", cada soldado se convierte en un nodo de un sistema de comunicaciones, y cada nodo busca y establece conexiones con nodos vecinos; todas estas conexiones, juntas, establecen los enlaces de la red. Cada nodo puede tanto enviar como recibir mensajes y servir de intermediario para retransmitir mensajes destinados a otros receptores. Así se extienden las capacidades de comunicación hasta mucho más allá de lo que puede transmitir cada nodo individual. Permite también una enorme versatilidad, porque la red viaja con sus usuarios, reconfigurando y restableciendo sin cesar sus conexiones según resulte necesario.

La codificación, al cambiar el funcionamiento de las redes, afectaría a la sociedad de maneras que todavía no podemos imaginar. En el ínterin, sin embargo, estamos estudiando los obstáculos que se oponen a su implantación. La conversión de nuestro sistema basado en enrutadores en otro basado en la codificación de red no sería más que un obstáculo poco importante. Tal conversión puede realizarse de forma gradual y no, necesariamente, de un golpe. Algunos enrutadores sólo deberán reprogramarse, mientras que otros, que no se construyeron para realizar codificaciones, se reemplazarían de una forma progresiva.

Un problema mayor será la resolución de las cuestiones que vayan surgiendo, más allá de la sustitución de los enrutadores por codificadores. Sea por caso la mezcla de informaciones: constituye una estrategia valiosa cuando el nodo receptor vaya a recibir indicios suficientes para recuperar lo que desea de entre aquella mezcla. Esta condición se cumple siempre en redes de multidifusión, pero podría no ser así en general. Además, en ciertas circunstancias, como ocurre cuando se emiten varias multidifusiones, la mezcla de información puede dificultar, si no impedir, que los usuarios recuperen lo que se querría recuperar. En tal caso, ¿de qué modo pueden los nodos decidir qué información puede mezclarse y cuál no cuando varias conexiones comparten una misma red? ¿En qué se ha de diferenciar la codificación de red que se aplique a los sistemas inalámbricos de la que se use en los tendidos fijos? ¿Qué ventajas ofrecería en lo que se refiere a la seguridad? ¿Cómo se habrá de facturar a los usuarios de servicios de comunicaciones, cuando los datos de una persona han de mezclarse con los de otros usuarios?

Los autores, y otros muchos investigadores, en colaboraciones que se extienden por todo el mundo, estamos reflexionando sobre la forma de resolver estas dudas, al mismo tiempo que nos esforzamos en ampliar la capacidad de las redes de comunicaciones, parte integral de muchas vidas.

Los autores

Michelle Effros, Ralf Koetter y **Muriel Médard** vienen colaborando desde hace tiempo. Effros enseña ingeniería eléctrica en el Instituto de Tecnología de California. Koetter dirige el Instituto de Ingeniería de Comunicaciones de la Universidad Técnica de Múnich. Médard es profesora de ingeniería eléctrica e informática en el Instituto de Tecnología de Massachusetts.

Bibliografía complementaria

A MATHEMATICAL THEORY OF COMMUNICATION. C. E. Shannon en *Bell System Technical Journal*, vol. 27, págs. 379-423 y 623-656; julio y octubre, 1948.

NETWORK INFORMATION FLOW. R. Ahlswede, N. Cai, S.-Y. R. Li y R. W. Yeung en *IEEE Transactions on Information Theory*, vol. 46, n.º 4, págs. 1204-1216; julio, 2000.

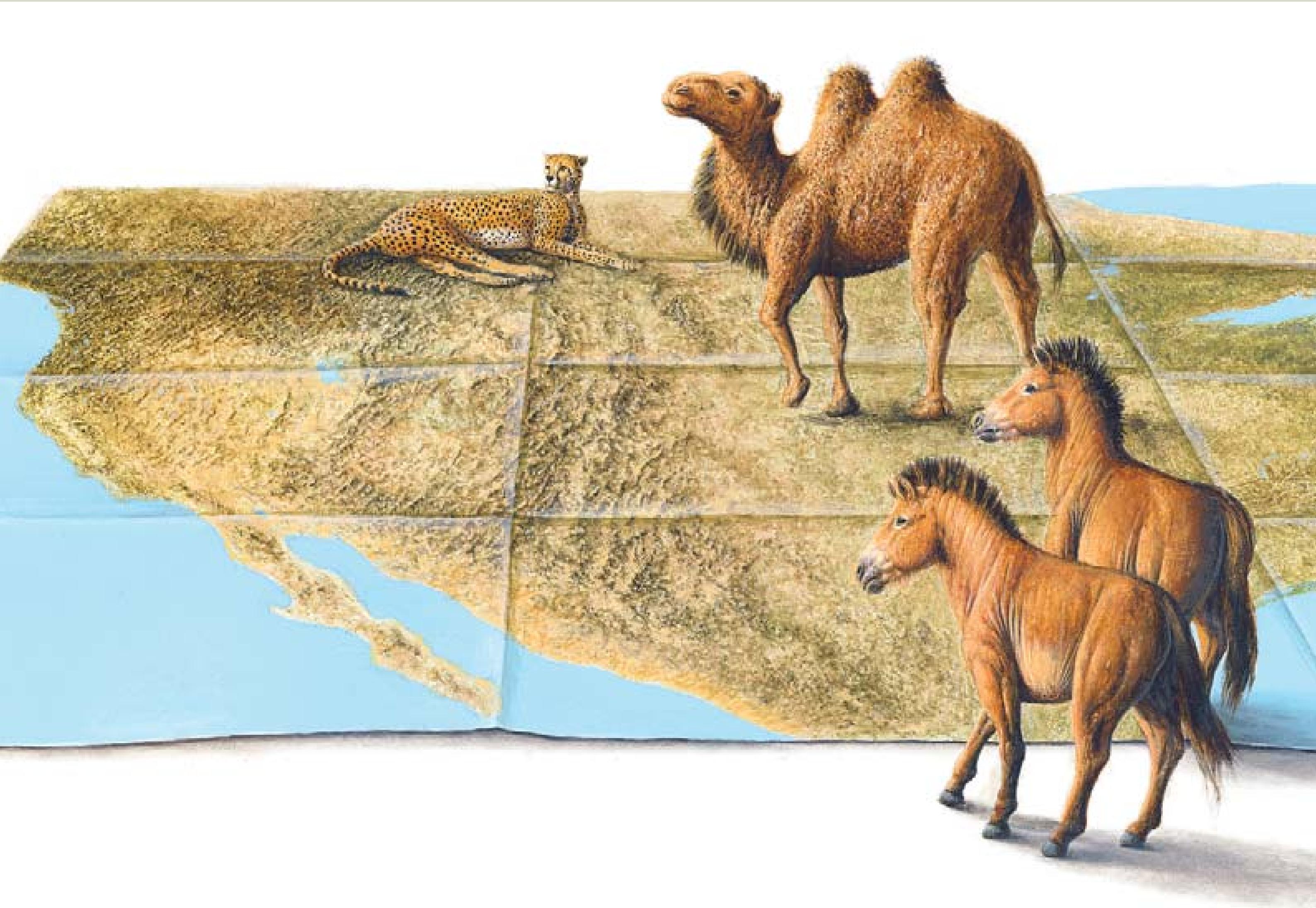
LINEAR NETWORK CODING. S.-Y. R. Li, R. W. Yeung y N. Cai en *IEEE Transactions on Information Theory*, vol. 49, n.º 2, págs. 371-381; febrero, 2003.

AN ALGEBRAIC APPROACH TO NETWORK CODING. R. Koetter y M. Médard en *IEEE/ACM Transactions on Networking*, vol. 11, n.º 5, págs. 782-795; octubre, 2003.

POLYNOMIAL TIME ALGORITHMS FOR MULTICAST NETWORK CODE CONSTRUCTION. S. Jaggi, P. Sanders, P. A. Chou, M. Effros, S. Egner, K. Jain y L. M. G. M. Toulhuizen en *IEEE Transactions on Information Theory*, vol. 51, n.º 6, págs. 1973-1982, junio, 2005.

A RANDOM LINEAR NETWORK CODING APPROACH TO MULTICASTING. T. Ho, M. Médard, R. Koetter, D. R. Karger, M. Effros, J. Shi y B. Leong en *IEEE Transactions on Information Theory*, vol. 52, n.º 10, págs. 4413-4430; octubre, 2006.

Recuperación de la



LARRY FELDER

En otoño de 2004, una docena de biólogos de la conservación estadounidenses nos reunimos en un rancho de Nuevo México para sopesar un plan ambicioso. En el grupo había expertos de varias disciplinas y de todas las edades, veteranos y noveles. La discusión se centró en la reintroducción de grandes vertebrados (megafauna) en los Estados Unidos.

La mayoría de esos animales, como los mamuts y guepardos, se extinguieron hace unos 13.000 años, cuando los humanos procedentes de Eurasia empezaron a colonizar el continente. Según la teoría que propuso hace 40 años Paul Martin, de la Universidad de Arizona, la sobrecaza por parte de los recién llegados causó una reducción tan drástica de las poblaciones de grandes vertebrados, que éstas no lograron recuperarse. La así llamada hipótesis de la "matanza del Pleistoceno" fue muy controvertida en su época. Pero hoy goza de amplia aceptación la idea de que los humanos tuvieron en ella un papel destacado.

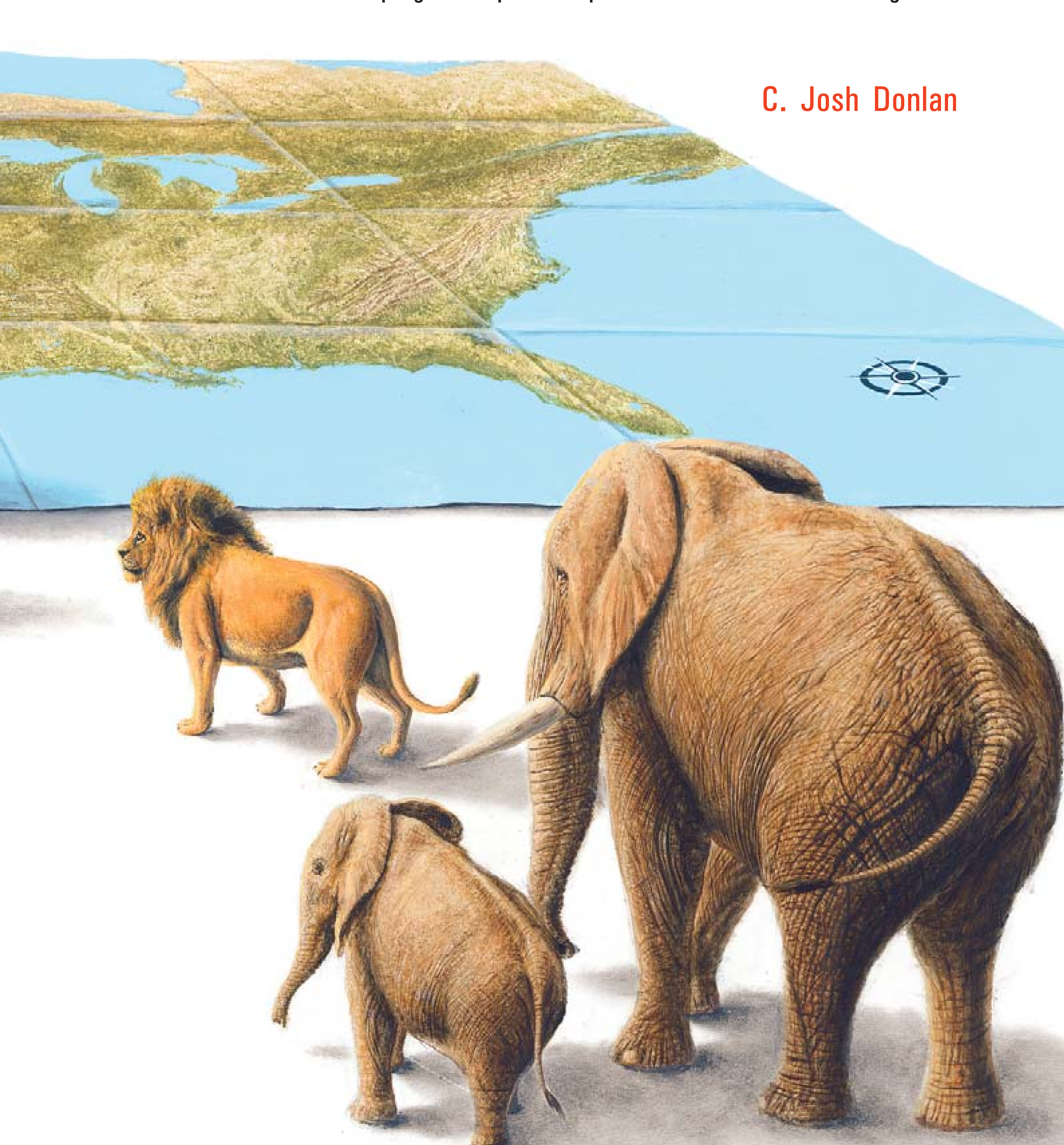
Martin estaba presente en la reunión de Nuevo México; sus ideas sobre la pérdida de esos animales, las consecuencias ecológicas y lo que deberíamos hacer a propósito de ello constituyeron los cimientos del proyecto que se forjó allí: la recuperación de la fauna pleistocénica.

Aunque extinguidos en Norteamérica los guepardos, leones y mamuts que antaño la recorrían a sus anchas, las mismas especies, o parientes cercanos, han sobrevivido en otros lugares. En las sesiones abordamos la introducción de esos sustitutos en ecosistemas estadounidenses. En nuestra opinión, ello permitiría restablecer de forma parcial procesos ecológicos de suma importancia, como la depredación y el ramoneo, en ecosistemas donde han estado ausentes desde hace milenios. La reintroducción de especies generaría también beneficios económicos y culturales. Como era de esperar, la publicación de nuestra propuesta provocó duras reacciones. Bienvenido sea el debate, pues los temas de conservación relacionados con la recuperación de la fauna pleistocénica merecen un examen exhaustivo.

fauna pleistocénica

Un proyecto para la recuperación de animales que desaparecieron de Norteamérica hace 13.000 años ofrece un programa optimista para la conservación en el siglo XXI

C. Josh Donlan



Importancia de la megafauna

Nuestro enfoque se centra en los animales de gran tamaño porque su efecto sobre el medio resulta de vital importancia. Durante decenas de millones de años, la megafauna dominó el planeta; su interacción y coevolución con otras especies influyó sobre ecosistemas enteros. Caballos, camellos, leones, elefantes y otros animales de porte vigoroso se encontraban por doquier: la megafauna era la norma. Pero desde hace unos 50.000 años, se extinguieron una abrumadora mayoría de esos animales. En la actualidad, la megafauna habita en una extensión inferior al 10 por ciento del globo.

En el curso de los diez últimos años, John Terborgh, ecólogo de la Universidad de Duke, ha comprobado en Venezuela la importancia de los animales de gran tamaño para la salud de los ecosistemas y el modo en que su pérdida perjudica al medio natural. Una presa hidroeléctrica venezolana inundó miles de hectáreas, con la aparición consiguiente de docenas de islas: una fragmentación del territorio semejante a la formación de islas virtuales que se produce con la tala de árboles, la construcción de centros comerciales, la expansión de centros urbanos y otras intervenciones humanas.

Pero las islas venezolanas, demasiado pequeñas, no podían abastecer a los animales instalados en la cúspide de las cadenas tróficas: jaguares, pumas, águilas y otros depredadores. Su desaparición desató una cadena de reacciones. Se multiplicaron los monos, las hormigas cortadoras de hojas y otros herbívoros cuya población había dejado de estar bajo el control de la depredación. Acabaron con la vegetación. Los ecosistemas se hundieron. La biodiversidad fue la gran perdedora.

Desastres ecológicos semejantes han acontecido en otros continentes. Los ecosistemas degradados no sólo son malos para la biodiversidad, sino que perjudican también a la economía. Se ha estimado que, en América Central, los ecosistemas tropicales suponen un valor de al menos 60.000 dólares anuales para la explotación cafetera, computados en razón de los servicios que aportan a la polinización de los cafetales.

Allí donde todavía quedan, depredadores y herbívoros de gran tamaño cumplen funciones cruciales. En Alaska, las nutrias marinas mantienen los ecosistemas de laminarias mediante el control de los herbívoros que comen ese tipo de algas, como los erizos de mar. En Africa, los elefantes configuran el paisaje: cuando se desplazan por el territorio, al derribar árboles y pisotear el suelo crean un hábitat en el que prosperan determinadas plantas y animales. Los leones y otros depredadores controlan las poblaciones de herbívoros africanos, que a su vez influyen sobre la distribución de la vegetación y los nutrientes del suelo.

Preguntas frecuentes

¿No morirán congelados los elefantes y leones en el invierno de Montana? ¿Se mantendrá a esos animales bajo techo, como se hace en los parques zoológicos?

La idoneidad del clima plantea una cuestión de interés para la recuperación de la fauna pleistocénica. Uno puede imaginar situaciones en las que los animales vivan en libertad (aunque con una gestión intensiva) durante todo el año en amplias reservas del sudoeste de los EE.UU.; otros escenarios se emplazarían en localidades más septentrionales, donde los animales se estabularían durante los meses fríos. Esta última opción es mejor que un parque zoológico. Ante las dificultades que entraña el cuidado de los elefantes (por sus necesidades espaciales), numerosos zoológicos prefieren descartarlos en sus programas.



ELEFANTE en el zoo de Copenhague.

¿Qué podemos aprender de los ranchos para la caza de animales salvajes que hay en Texas?

Existen numerosas fincas dedicadas a la cría de animales de caza mayor en el estado de Texas; en algunas propiedades hay animales que, según indica el registro fósil, vivieron antaño en Norteamérica (guepardos, por ejemplo). Hasta la fecha, ningún biólogo de la conservación ha estudiado los animales que se crían en tales explotaciones ganaderas, pero los propietarios podrían ofrecer un servicio excelente a la investigación con su disponibilidad.

¿Se había planteado antes una propuesta semejante?

La idea de recuperar la fauna pleistocénica surgió hace ya algún tiempo. La reintroducción corresponde a la práctica de volver a implantar especies en lugares de donde desaparecieron en el curso de los últimos siglos. La recuperación de la fauna pleistocénica, en cambio, implica la introducción de especies descendientes (o parientes próximos) de animales que se extinguieron hace unos 13.000 años.

Resumen/*Recuperación de la fauna pleistocénica*

- Un grupo conservacionista ha propuesto la reintroducción en EE.UU. de animales de gran tamaño que se extinguieron hace 13.000 años.
- Parientes cercanos de esos animales —elefantes, camellos, leones o guepardos— sobrevivieron en otras partes del planeta; su reintroducción en EE.UU. restablecería procesos ecológicos clave que antaño se daban allí, al propio tiempo que proporcionaría un refugio a especies amenazadas de Africa y Asia, y crearía oportunidades de negocio para el ecoturismo.
- El proyecto tiene también sus detractores. Con todo, tan atrevida propuesta ha dado un giro positivo a la biología de la conservación, que hasta ahora se había centrado en la lucha contra la pérdida de biodiversidad.

DONA *Peter Arnold, Inc.*

En la América del Pleistoceno, los depredadores y herbívoros de gran tamaño desempeñaban funciones similares. En la actualidad, su influencia ecológica brilla por su ausencia. Sea por caso el guepardo americano (pariente del guepardo africano). Durante millones de años, recorrió las praderas a la caza de berrendos, o antílopes americanos, persecuciones que modelaron la locomoción y otros aspectos fisiológicos del berrendo, convirtiéndolo en uno de los animales más veloces. En ausencia del guepardo, el berrendo parece "dotado en demasía" para su medio actual.

La recuperación de la fauna pleistocénica no busca una recreación exacta del pasado. Se propone restablecer las interacciones entre especies que garantizan la prosperidad de los ecosistemas. Tortugas gigantes, caballos, camellos, guepardos, elefantes y leones: todos estaban ahí y ayudaron a modelar los ecosistemas norteamericanos. Se trataría de introducir esas especies, u otras estrechamente emparentadas. Un buen número de esas especies se encuentra ya en cautividad en los Estados Unidos. En esencia, la repoblación pleistocénica contribuiría a invertir la premisa básica de la biología de la conservación: se pasaría de frenar la extinción a la restauración de procesos naturales.

De entrada, nuestra propuesta puede parecer extravagante. ¿Leones en Montana? Pero el plan merece un debate serio por varias razones. Primera, ya no quedan en la Tierra lugares vírgenes, al menos en el sentido de estar libres de la influencia humana. Nuestra demografía, nuestros productos químicos, nuestra economía y política impregnan todos y cada uno de los rincones del planeta. Incluso en los mayores parques nacionales, las especies se extinguen sin intervención activa. Y la penetración del hombre en las áreas naturales muestra síntomas alarmantes de empeoramiento. La inversión de esa tendencia negativa requiere la aplicación de medidas nuevas y más arriesgadas que las que se llevan a cabo en la actualidad.

En segundo lugar, desde su nacimiento hace más de tres decenios, la historia de la biología de la conservación ha consistido en una sucesión de catástrofes y desalientos, en una

Un Oeste salvaje

POCO DESPUES de que los humanos cruzaran el puente continental de Bering y penetraran en Norteamérica, hace unos 13.000 años, casi el 75 por ciento de los mamíferos de gran tamaño del continente (los que pesaban más de 45 kilogramos) desaparecieron (*rojo*). Uno de los objetivos del proyecto de recuperación de la fauna pleistocénica es restablecer algunas especies, o sus parientes próximos, en el oeste estadounidense. El león y el guepardo, comunes otrora en Norteamérica, perviven ahora sólo en Africa. El elefante africano o asiático sustituiría al extinguido mamut; los camellos bactrianos podrían reemplazar al desaparecido *Camelops*.

Grandes mamíferos del Pleistoceno de Norteamérica

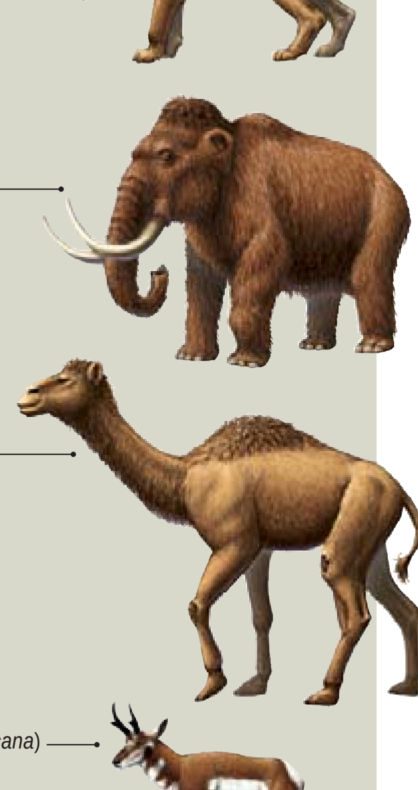
Xenartros
- **Gliptodonte** (*Glyptotherium floridanum*)
- **Perezoso terrestre de Harlan** (*Paramylodon harlani*)
- **Perezoso terrestre de Jefferson** (*Megalonyx jeffersonii*)
- **Perezoso terrestre de Shasta** (*Nothrotheriops shastensis*)

Carnívoros
- **Lobo terrible** (*Canis dirus*)
- **Lobo gris** (*Canis lupus*)
- **Oso negro** (*Ursus americanus*)
- **Oso pardo** (*Ursus arctos*)
- **Oso caricorto gigante** (*Arctodus simus*)
- **León de dientes de sable** (*Smilodon fatalis*)
- **León americano** (*Panthera leo*)
- **Jaguar** (*Panthera onca*)
- **Guepardo americano** (*Miracinonyx trumani*)
- **Puma** (*Puma concolor*)

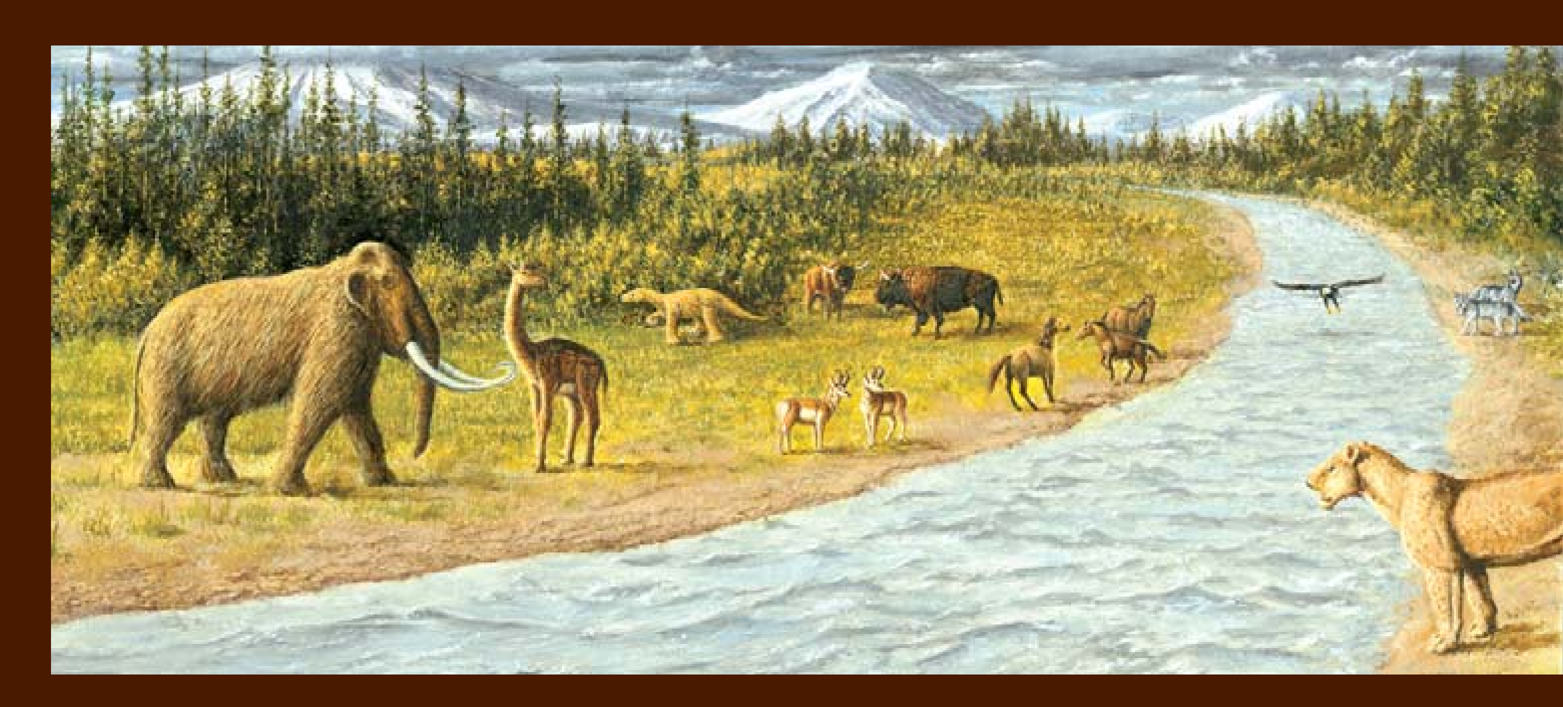
Elefantes (Proboscídeos)
- **Mastodonte americano** (*Mammut americanum*)
- **Mamut colombino** (*Mammuthus columbi*)
- **Mamut enano** (*Mammuthus exilis*)
- **Mamut lanudo** (*Mammuthus primigenius*)

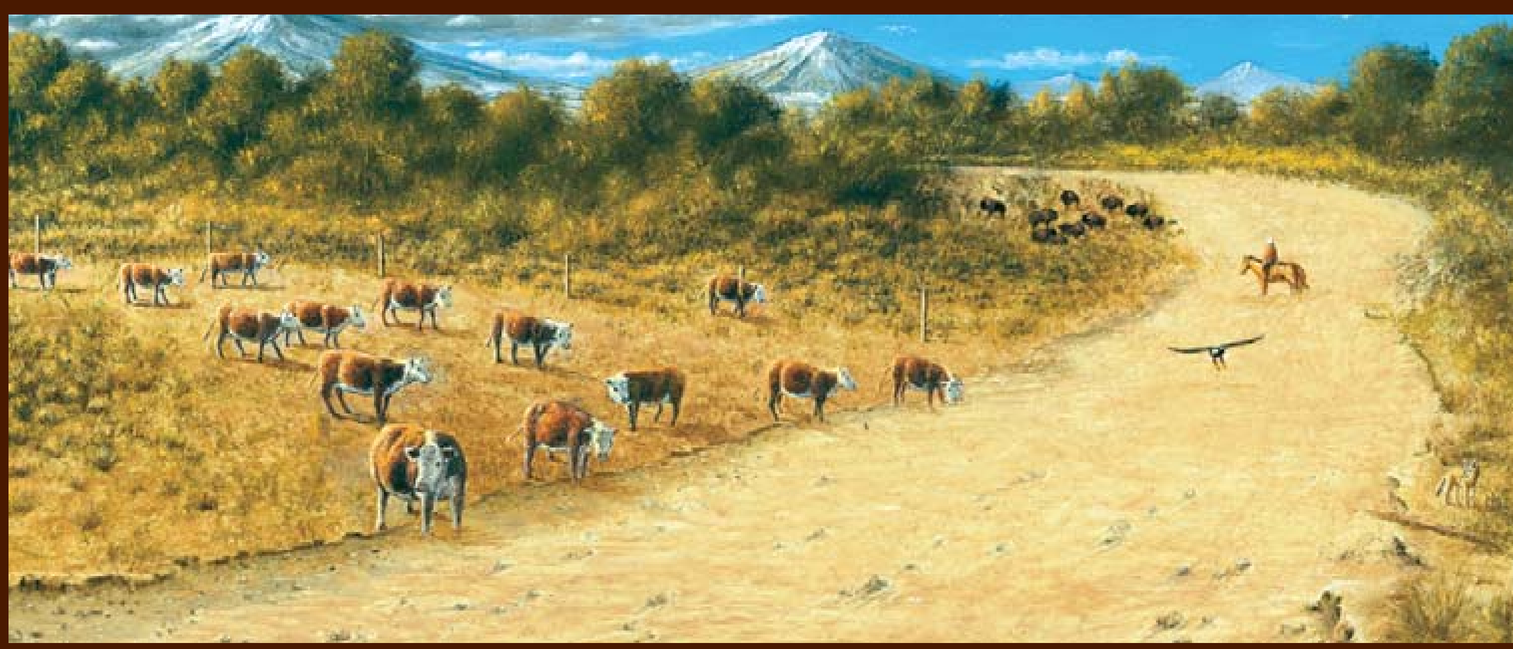
Caballos (Perisodáctilos)
- **Caballo mexicano** (*Equus conversidens*)
- **Caballo occidental** (*Equus occidentalis*)
- **Otros caballos y asnos extinguidos** (*Equus spp.*)

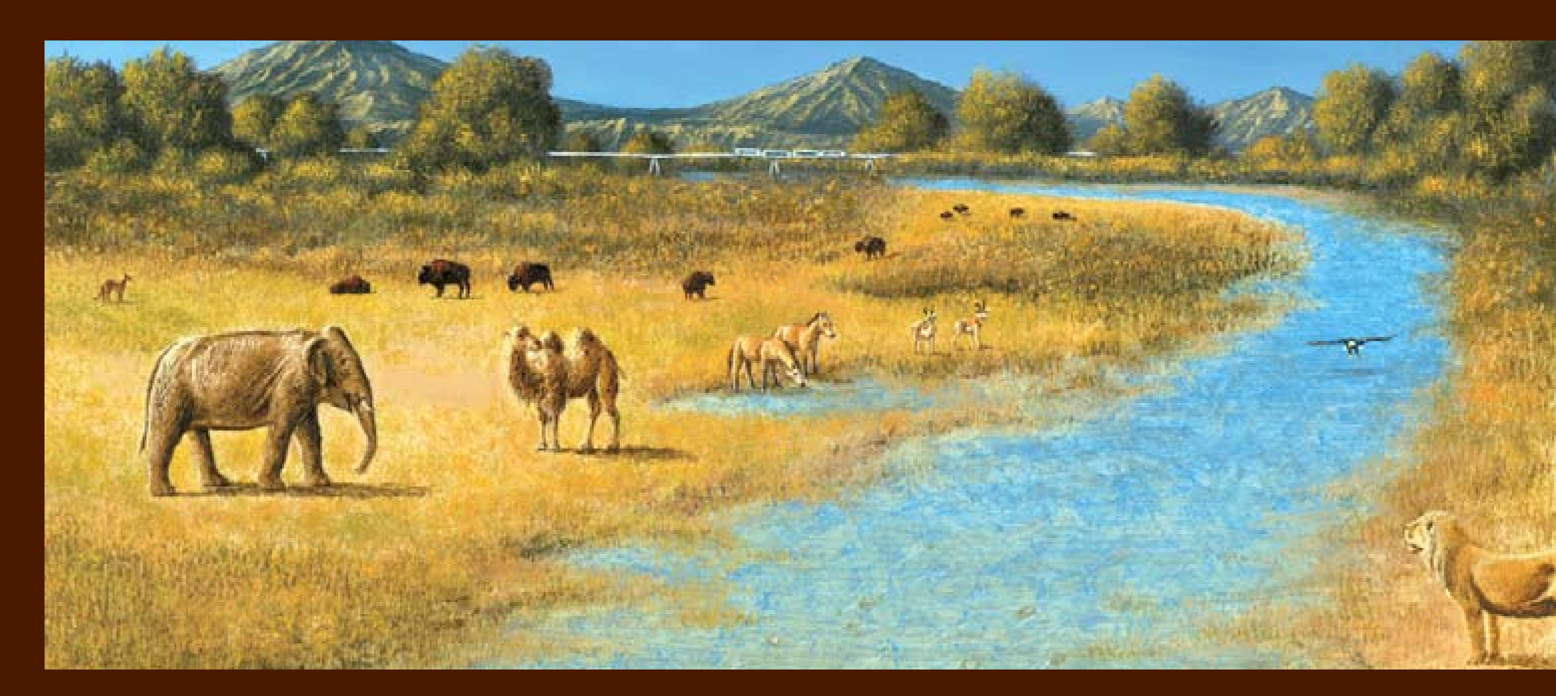
Ungulados de dedos pares (Artiodáctilos)
- **Camello occidental** (*Camelops hesternus*)
- **Llama patilarga** (*Hemiauchenia macrocephala*)
- **Pécari narigudo** (*Mylohyus nasutus*)
- **Pécari de cabeza plana** (*Platygonus compressus*)
- **Ciervo mula** (*Odocoileus hemionus*)
- **Ciervo de Virginia** (*Odocoileus virginianus*)
- **Ciervo de montaña** (*Navahoceros fricki*)
- **Caribú** (*Rangifer tarandus*)
- **Alce** (*Alces alces*)
- **Uapití** (*Cervus elaphus*)
- **Berrendo o antílope americano** (*Antilocapra americana*)
- **Cabra blanca de montaña de Harrington** (*Oreamnos harringtoni*)
- **Cabra blanca de montaña** (*Oreamnos americanus*)
- **Muflón canadiense** (*Ovis canadensis*)
- **Buey de matorral** (*Euceratherium collinum*)
- **Buey almizclero de bonete** (*Bootherium bombifrons*)
- **Bisonte** (*Bison bison*)
- **Bisontes extinguidos** (*Bison spp.*)



SHAWN GOULD







Pasado, presente y futuro

Estas reproducciones artísticas ilustran algunos de los efectos que los animales de gran tamaño ejercen sobre el ecosistema y los beneficios que aportaría la repoblación pleistocénica. Las ilustraciones son fruto de la colaboración entre David Burney, del Jardín Botánico Tropical Nacional de EE.UU., miembro del grupo de recuperación de la fauna pleistocénica, y Larry Felder, paleoartista.

◄◄ **FINALES DE ABRIL, DURANTE EL PLEISTOCENO, HACE 14.000 AÑOS**
Los humanos no han aparecido todavía en el paisaje. Los glaciares se han retirado; el clima de lo que hoy son los EE.UU. occidentales es parecido al actual, aunque un poco más frío. Las Montañas Rocosas están cubiertas por nieve densa. El río, alimentado por la fusión de las nieves, baja alto; su color azul lechoso se debe a las partículas finas que transporta, procedentes de los suelos jóvenes posglaciales. Las orillas, pastadas intensamente, suelen mantenerse verdes. Un mamut colombino, un *Camelops*, un perezoso terrestre de Shasta, bisontes cornilargos, caballos y antílopes berrendos están pastando, mientras un león americano y dos lobos los observan desde la orilla opuesta. Un águila calva planea sobre el río.

◄◄ **ABRIL DE 2007**
Las Montañas Rocosas tienen nieve sólo en sus cumbres más altas. El río, cenagoso por las lluvias de abril y por un exceso de ganado en las orillas, se ha ensanchado; las orillas se han erosionado y desmoronado debido al pisoteo y a los revolcaderos del ganado. La hierba es corta, hay retazos de suelo denudado y predominan los matorrales espinosos. A lo lejos, un rebaño de bisontes pasta, mientras un coyote aguarda cauteloso en la orilla opuesta.

◄◄ **ABRIL DE 2027**
La vegetación es más estival: las condiciones son de caldeamiento global. Ya no hay nieve en las cumbres alpinas. La sequía, la baja humedad y la ausencia de nieve en las montañas han hecho descender de forma espectacular el nivel del río; el agua es transparente y oscura. La erosión, menos manifiesta; la hierba está recortada uniformemente por el pastoreo y el ramoneo. Un elefante asiático, un camello bactriano, bisontes, antílopes y caballos de Przewalski pastan bajo la mirada escrutadora de un león africano; desde la distancia se acerca un guepardo. Al fondo, un monorraíl transporta turistas. Una valla electrificada de alta tecnología, con paneles solares, queda fuera de la vista.

LARRY FELDER

lucha por frenar la pérdida de biodiversidad. Pero la conservación no debe operar sólo a la contra. Nosotros proponemos un enfoque activo, que contemple el restablecimiento de los procesos naturales, empezando por los que revisten mayor importancia, como los que dependen de la presencia de megafauna.

En tercer lugar, en EE.UU. disponemos de tierras para la reintroducción de megafauna. Aunque los patrones de uso humano del suelo son siempre cambiantes, podrían utilizarse para el proyecto algunas de las amplias extensiones que cubren las Grandes Llanuras y el sudoeste estadounidenses. Pudieran ser de terrenos de titularidad pública o privada, con densidad de población humana reducida o menguante.

En cuarto lugar, la presencia de megafauna pleistocénica favorecería el turismo y otras actividades económicas en las comunidades vecinas, al tiempo que aumentaría el interés del público por la naturaleza. Más de 1,5 millones de personas visitan cada año el Parque de Fieras de San Diego para ver, aunque sólo sea fugazmente, mamíferos de gran tamaño. Pocos parques nacionales estadounidenses reciben tantos visitantes. Por fin, parece probable que se produzca, en el transcurso de este siglo, la pérdida de algunas de las especies de megafauna que quedan en Africa y Asia; la recuperación de la fauna pleistocénica ayudaría a invertir la tendencia.

Fases del proyecto

No se trata de tomar un camión y descargar algunos guepardos en el jardín más cercano. Tampoco estamos hablando de hacerlo mañana. La repoblación pleistocénica debe entenderse en el marco de una serie de medidas en el ecosistema cuidadosamente preparadas y controladas. Ofrecemos aquí algunas recomendaciones —no un programa— de los pasos a dar, sin que pretendamos anteponer nuestra propuesta a los programas de conservación vigentes, de EE.UU. o Africa. La recuperación de la fauna pleistocénica se realizaría a la par que esos proyectos de conservación; se obtendrían fondos procedentes de nuevas fuentes de financiación, sin competir, pues, por los fondos de los programas de conservación que ya existen.

La última fase del proyecto consistiría en la construcción de un parque ecológico histórico, donde vagarían caballos, camellos, elefantes y carnívoros de gran tamaño; vallado por mor de la seguridad, cubriría una extensión de miles de kilómetros cuadrados. Tal como ocurre hoy en Africa y en las regiones que rodean algunos parques nacionales estadounidenses, el parque ecológico histórico no sólo atraería ecoturistas, sino que crearía también puestos de trabajo relacionados con la gestión del parque y el turismo.

Las primeras acciones, en cambio, serían modestas: estudios preliminares a escala reducida que evaluarían el impacto de la megafauna en los paisajes estadounidenses. Tales experimentos controlados se basarían en el registro fósil (que indica qué animales vivieron en una región determinada); se llevarían a cabo en terrenos privados, donados o comprados. Podrían empezar de inmediato. Darían respuesta a numerosas preguntas acerca de las reintroducciones, además de facilitar la evaluación de los costes y los beneficios de la repoblación.

Ha empezado a andar ya el primer experimento. Animados por nuestra reunión de 2004, los biólogos reintrodujeron tortugas llaneras o del Bolsón (*Gopherus flavomarginatus*) en un rancho privado de Nuevo México. Las tortugas, propias del Bolsón de Mapimí, algunas de las cuales pesan más de 40 kilogramos, ramoneaban en regiones del sudoeste de los EE.UU., antes de desaparecer hace 10.000 años, víctimas de la caza humana. Esa tortuga amenazada se aferra ahora a la supervivencia, acantonada en una área exigua del centro de Nuevo México. De ese modo, la reintroducción no sólo repatría la tortuga a los EE.UU., sino que, además, aumenta las probabilidades de supervivencia de la especie. Experimentos similares se están llevando a cabo en otros países.

Otros de los protagonistas de los experimentos iniciales de reintroducción serían los caballos salvajes y los camellos, ambos originarios de América. En el Pleistoceno tardío abundaban las especies de esos mamíferos. Los caballos y asnos cimarrones que en la actualidad viven en algunas áreas del oeste estadounidense serían unos buenos sustitutos de las especies americanas extinguidas. Puesto que la

mayoría de las especies eurasiáticas y africanas supervivientes se hallan ahora en peligro crítico de extinción [*véase* “Equidos salvajes en peligro”, de Patricia D. Moehlman; INVESTIGACIÓN Y CIENCIA, mayo de 2005], el establecimiento en Norteamérica de asnos asiáticos y el caballo de Przewalski podría ayudar a evitar la extinción de estos animales. Los camellos bactrianos, amenazados de extinción en el desierto de Gobi, proporcionarían un sustituto moderno para *Camelops*, un camello del Pleistoceno tardío. Los camellos, introducidos a partir de poblaciones en cautividad o domesticadas, podrían enriquecer a los ecosistemas estadounidenses, puesto que ramonean plantas leñosas que se expanden por las praderas áridas del sudoeste, un ecosistema en riesgo creciente.

Otra fase del proyecto, más controvertida, podría empezar también de inmediato. Se trata del establecimiento en propiedades privadas de un número reducido de elefantes, guepardos y leones.

De la introducción de elefantes se beneficiarían también las poblaciones humanas cercanas, al convertirse en una atracción turística, y los ganaderos, al preservar las praderas; los elefantes suprimirían las plantas leñosas que amenazan las praderas sudoccidentales. En el Pleistoceno tardío, medraban en Norteamérica al menos cuatro especies de elefantes. Los ejemplares que hoy viven en cautividad en EE.UU. se introducirían en sustitución de sus homólogos extinguidos. El mayor coste correspondería a la construcción de vallas, que han contribuido a reducir los conflictos entre elefantes y humanos en Africa.

En cuanto a los guepardos, ya viven en cautividad un buen número de ellos. El mayor reto sería proporcionarles áreas cercadas y extensas, que incluyeran los hábitats y las presas adecuadas. Los costes de la operación se compensarían con el restablecimiento de interacciones fuertes con los berrendos, el fomento del ecoturismo y el freno de la extinción del carnívoro más veloz del mundo [*véase* “El guepardo, en peligro de extinción genética”, de Stephen J. O’Brian, David E. Wildt y Mitchell Bush; INVESTIGACIÓN Y CIENCIA, julio de 1986].

Los leones viven bajo una amenaza creciente; hay poblaciones en Asia y algunas regiones de Africa en peligro crítico de extinción. La recuperación del león, la misma especie que antaño vivió en Norteamérica, presenta desafíos imponentes, a la vez que beneficios potenciales. Pero las reservas privadas del Africa austral en las que se han reintroducido con éxito leones y otros animales de gran tamaño ofrecen un buen ejemplo a seguir; las reservas ocupan allí extensiones menores que algunos ranchos privados del sudoeste estadounidense.

Si esos primeros experimentos con herbívoros y depredadores poderosos muestran resultados prometedores, el proyecto seguiría adelante, hacia la construcción final de un extenso parque ecológico histórico. Lo que necesitamos ahora son paneles de expertos que, para cada especie, asesoren, evalúen y dirijan con prudencia el restablecimiento de la megafauna en Norteamérica.

Un ejemplo real de reintroducción de un depredador culminante procede de los lobos del Parque Nacional de Yellowstone. El lobo

La opinión de los críticos

DESDE QUE EN 2005 propusimos por escrito la recuperación de la fauna pleistocénica, en la revista *Nature*, se han esgrimido varias objeciones. Unas son legítimas, otras no. Según escribieron Dustin Rubenstein y sus colaboradores en la revista *Biological Conservation*, la recuperación de la fauna pleistocénica de Norteamérica es una propuesta poco menos sensacionalista que el fantasioso *Parque Jurásico*. No estamos de acuerdo con esa afirmación. La mayoría de los dinosaurios se extinguieron hace 65 millones de años, mientras que buena parte de la megafauna originaria de Norteamérica, o parientes muy próximos, perviven en otros lugares del mundo, lo que permite estudiarlos y salvarlos de la extinción.
Rubenstein aduce que las especies equivalentes modernas son “incorrectas”, genéticamente distintas de las especies que medraron en Norteamérica durante el Pleistoceno. Cierto, aunque no tan distintas. Las pruebas indican que los leones actuales de Africa y Asia pertenecen a la misma especie, aunque son de tamaño menor, que los leones que merodeaban por Norteamérica hace 13 milenios. Estudios recientes de ADN antiguo han dilucidado relaciones de parentesco próximas entre las especies de elefantes y caballos extinguidas y las que viven en la actualidad. Además, la introducción de especies equivalentes a otras extinguidas ya se ha realizado con éxito en otros experimentos. Por citar un caso, la utilización de cientos de halcones peregrinos de Australia, Europa y Sudamérica en programas de reproducción en cautividad para la reintroducción del halcón peregrino en zonas de los EE.UU. y Canadá donde el DDT los había exterminado. Esas aves eran, ciertamente, distintas desde el punto de vista genético de las que antaño se cernían sobre el Medio Oeste, pero se han adaptado sin mayores problemas a su nuevo hogar.
Rubenstein pronostica un escenario hostil en el que los granjeros deberán enfrentarse a la destrucción de las cosechas por parte de rebaños de elefantes, mientras el ganado, e incluso los niños, serán víctimas del ataque de leones y guepardos. Pero no hace falta imaginar los retos que conlleva la coexistencia con depredadores y herbívoros de gran tamaño. Africa y Asia han estado bregando con ellos durante siglos y se ha logrado un progreso notable. Además, en nuestro plan los animales no vivirían en completa libertad.
Para Christopher Smith, el cambio climático global acontecido desde las extinciones del Pleistoceno hace que el restablecimiento de ecosistemas desaparecidos mediante la introducción de mamíferos de gran tamaño sea muy improbable. Muchos expresaron su preocupación por el hecho de que los ecosistemas de Norteamérica no son los mismos que hace 13.000 años, lo que dificultaría la supervivencia de los animales introducidos. Aunque los hábitats son y seguirán siendo dinámicos a una escala temporal de miles de años, muy pocas plantas o micromamíferos se extinguieron durante el Pleistoceno tardío. El gran ausente de los ecosistemas norteamericanos actuales es la megafauna, de lo que se infiere que constituye un elemento clave para la salud de los mismos.

TINA CHOWDHURY

Reintroducción de mamíferos en España

LA POLEMICA PROPUESTA del grupo de Josh Donlan, sobre la reconstrucción de fauna pleistocénica en EE.UU. como medida activa de conservación que permita restaurar procesos ecológicos como la depredación y el ramoneo, debería valorarse también para otras áreas geográficas como Europa. Nuestro caso difiere sustancialmente del ejemplo norteamericano. El territorio europeo, muy humanizado, alberga una superficie mucho menor de espacios naturales. Unos espacios que, en realidad, son poco "naturales", pues en su mayoría han sido modificados por la actividad humana, difieren de los ecosistemas primigenios y carecen de funciones clave como las que ejercían los grandes depredadores.

Por otra parte, en Europa occidental el bioma natural dominante es el bosque. Las extensas praderas (como las de Norteamérica) no suelen formar parte de la vegetación autóctona. Los herbívoros de gran tamaño desempeñan una función clave en el mantenimiento de las praderas, pero no en el mantenimiento del bosque; más bien al contrario, pues impiden la regeneración del mismo.



Lobo europeo (*Canis lupus lupus*)

Por albergar una diversidad extraordinaria de ecosistemas, flora y fauna, la península Ibérica constituye una excepción en Europa. Además, aunque aisladas y, en ocasiones, amenazadas, mantiene poblaciones de depredadores de gran tamaño que han desaparecido de la mayor parte de Europa: el lobo, el oso y el lince. Además, esos depredadores corresponden a las mismas especies, o a descendientes directos, que dominaban la península Ibérica hace miles de años. Por tanto, bien podrían desempeñar la función señalada por Josh Donlan.

La península Ibérica cuenta con una larga lista de introducciones y reintroducciones de herbívoros de gran tamaño, casi siempre con fines cinegéticos. Ciervos y cabras monteses se han reintroducido en numerosos territorios en los que habían desaparecido. Gamos, muflones y arruís son especies alóctonas (no pertenecientes a la fauna ibérica) características de zonas áridas que se han adaptado con éxito en algunas sierras de la península Ibérica. Aunque no puede considerarse la única causa, la ausencia de depredadores y, por tanto, de su función reguladora, ha favorecido el crecimiento desproporcionado de las poblaciones de esos herbívoros; la regulación natural se ha producido por la acción de los parásitos (por ejemplo, el efecto de la sarna sobre la cabra montés en la sierra de Cazorla).

Aunque el caso del oso es especial —por tratarse de un animal fundamentalmente frugívoro y carroñero, ejerce un control escaso sobre las poblaciones de herbívoros de gran tamaño—, ilustra la problemática derivada de su posible reintroducción. Los intentos que se han llevado a cabo para reforzar la maltrecha población pirenaica de osos han topado con una fuerte oposición de las poblaciones locales, francesas y españolas.

El lobo constituye el más abundante de los grandes depredadores y el que ejerce mayor control sobre los ungulados, ya que forman parte de su dieta. La gestión de esa especie figura entre los retos de mayor complejidad que afrontan los responsables del medio natural. Cada año se matan en la península Ibérica varios centenares de lobos; la mayoría de las veces de forma ilegal; un buen número de ellos, en espacios naturales protegidos. Su reintroducción para restaurar su función ecológica de depredador no se aceptaría fácilmente; así ha ocurrido en EE.UU., donde se han reintroducido lobos en algunos estados.

Menos dificultades entrañaría la reintroducción del lince ibérico, uno de los felinos más amenazados del mundo y especializado en la depredación de conejos. Solventados los problemas técnicos que conlleva la cría en cautividad y la suelta en el medio natural, ésta no tendría efectos contra los intereses de la población.

Si parte de la opinión pública no está a favor de la reintroducción de carnívoros con los que ha convivido recientemente y se convive todavía en algunos territorios (osos y lobos, por ejemplo), es poco probable que viera con buenos ojos un programa que proyecte la restauración de depredadores pleistocénicos (leones, panteras y hienas).

En la parte positiva cabe indicar que no sería necesaria la reconstrucción de la fauna de herbívoros de gran tamaño extinguidos en los últimos milenios (elefantes, rinocerontes, bisontes, uros) para mantener los ecosistemas pastados y ramoneados. Los herbívoros actuales —verdaderos "ingenieros ecosistémicos"— han moldeado el paisaje que hoy se encuentra sobrepastado, sobre todo en áreas mediterráneas donde, en una auténtica sinergia con la sequía, comprometen seriamente la regeneración natural del bosque y el matorral mediterráneo. En esas zonas sí parecería pertinente el restablecimiento de depredadores de gran tamaño (reintroducción o reforzamiento de poblaciones de lobos) para restaurar las funciones del ecosistema.

José Ramón Obeso
Universidad de Oviedo

gris se extinguió en Yellowstone y sus aledaños en los años veinte del siglo pasado. Esa pérdida condujo a un aumento de las presas (alces y ciervos) que, a su vez, redujo la población de álamos temblones y de otros árboles que integran su dieta. La falta de vegetación destruyó el hábitat para las aves migratorias y los castores. Así, la desaparición de los lobos propagó una cascada trófica, de los depredadores a los herbívoros, de éstos a las plantas, a las aves y a los castores. Los expertos han empezado a documentar los cambios operados en el ecosistema a medida que los lobos reintroducidos recuperan la función ecológica que desempeñaron en Yellowstone durante milenios.

La reintroducción de lobos en Yellowstone ha demostrado que éstos podrían ayudar al parque a mitigar los efectos del cambio climático. Con unos inviernos más suaves, mueren menos ciervos, es decir, menos alimento para coyotes, cuervos, águilas calvas y otros carroñeros. Los lobos proporcionan cadáveres durante todo el año para que los carroñeros se alimenten, lo que confiere cierta estabilidad ecológica.

Problemas que deben abordarse

Con todo, la recuperación de la fauna pleistocénica entraña numerosos obstáculos. Entre éstos, la posibilidad de que los animales introducidos aporten enfermedades nuevas o que se hallen expuestos a enfermedades preexistentes en el ecosistema. Al haber cambiado tanto los hábitats en el transcurso de los milenios, podría ocurrir que los animales introducidos no se adaptaran a un medio alterado. Sin olvidar considerar la probabilidad de consecuencias ecológicas imprevistas y de reacciones inesperadas de las comunidades humanas vecinas.

En otras partes del mundo

En otras regiones del mundo se han iniciado programas preliminares para la reintroducción de especies allí donde se extinguieron.

- En abril de 2006, un equipo de biólogos canadienses y rusos transportaron en avión 30 bisontes de bosque desde el Parque Nacional de la Isla Elk, de Canadá, hasta la reserva del Parque del Pleistoceno, en la república de Sakha, en Rusia, de donde el bisonte de estepa, estrechamente emparentado, desapareció hace 5000 años.
- En la reserva holandesa de Oostvaardersplassen, de 6000 hectáreas, los conservacionistas están reintroduciendo caballos, corzos y uros.
- En toda Europa se está llevando a cabo la reintroducción de castores, en algunos casos en regiones donde estaban ausentes desde hace miles de años.
- En el Pacífico tropical, aves en peligro de extinción de las islas Marquesas y Tonga se han reintroducido en islas cercanas donde, según el registro fósil, vivían antaño.
- En el océano Indico, científicos de la Fundación de Vida Salvaje de Mauricio utilizan tortugas gigantes de la isla de Aldabra para sustituir a dos especies de tortugas extinguidas de las islas Mascareñas. Ya se ha registrado un aumento de la dispersión de semillas para numerosas especies de plantas insulares. La presencia de tortugas también ha provocado un aumento del turismo.



CABALLOS KONIK ocupan el lugar del extinguido tarpán en la reserva holandesa de Oostvaardersplassen.

STEVEN RUITER *Foto Natura/Minden Pictures*

El autor

C. Josh Donlan se doctoró en ecología y biología evolutiva por la Universidad de Cornell, donde hoy lleva a cabo su labor investigadora. Fundador y director de Estrategias Avanzadas de Conservación, es también asesor del Parque Nacional de Galápagos y de la organización ecologista para la Conservación de las Islas. Pasa largas estancias en Tasmania, Australia y Santa Cruz, ocupado en frenar las extinciones insulares.

Bibliografía complementaria

THE GHOSTS OF EVOLUTION: NONSENSICAL FRUIT, MISSING PARTNERS, AND OTHER ECOLOGICAL ANACHRONISMS. Connie Barlow. Basic Books, 2000.

TWILIGHT OF THE MAMMOTHS: ICE AGE EXTINCTIONS AND THE REWILDING OF AMERICA. Paul S. Martin. University of California Press, 2005.

PLEISTOCENE REWILDING: AN OPTIMIST AGENDA FOR TWENTY-FIRST CENTURY CONSERVATION. C. J. Donlan, J. Berger, C. E. Bock, J. H. Bock, D. A. Burney, J. A. Estes, D. Foreman, P. S. Martin, G. W. Roemer, F. A. Smith, M. E. Soulé y H. W. Greene en *American Naturalist*, vol. 168, n.º 5, págs. 660-681; noviembre, 2006.

El desarrollo de programas que se encarguen del seguimiento de las interacciones entre especies y sus consecuencias ecológicas requerirá paciencia y experiencia. Y, desde luego, no será fácil convencer a la opinión pública de que la depredación constituye un proceso natural de importancia, que nutre la tierra y favorece la prosperidad de los ecosistemas.

Muchos pensarán que la idea de la recuperación de la fauna pleistocénica es inviable en nuestros días. Pero dirijamos la mirada a Africa. Transcurrido un año desde la creación del Parque Nacional Kruger, se hallaba éste muy lejos de constituir el pilar actual de la biodiversidad del Africa austral. En 1903, vivían dentro de sus límites nueve leones, ocho búfalos y escasos guepardos; ningún elefante. Merced a la labor de los conservacionistas africanos, 100 años más tarde vagaban por el Kruger 7300 elefantes, 2300 leones, 28.000 búfalos y 250 guepardos... al igual que lo hacían 700.000 turistas, que se dejaban decenas de millones de dólares.

En la centuria recién iniciada, la humanidad decidirá, por acción u omisión, la medida en que toleraremos a otras especies y, por tanto, cuánta diversidad perdurará. La recuperación de la fauna del Pleistoceno no tiene nada que ver con una vuelta al pasado; se trata de recuperar del pasado estrategias para el mantenimiento de la estructura funcional de la naturaleza. Los beneficios científicos, de conservación y culturales que generaría el restablecimiento de la megafauna son claros, como lo son los costes. Aunque una planificación científica rigurosa mitigaría los costes potenciales, la propuesta inquietaría a más de uno. Pero, dada la disfunción aparente de los ecosistemas de Norteamérica y del estado global de la Tierra, la actitud pasiva entraña sus propios riesgos.

Ante un futuro de incertidumbres, los expertos y la sociedad deben sopesar los costes y los beneficios que comportarían acciones atrevidas y contundentes como la recuperación de la fauna pleistocénica, comparados con los riesgos, incertidumbres y costes de la inactividad y que no suelen reconocerse. Tendemos a pensar que si mantenemos el *status quo*, todo irá bien. Sin embargo, los datos disponibles apuntan lo contrario.

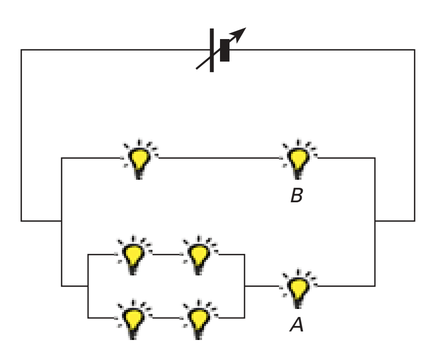
CURIOSIDADES DE LA FÍSICA

Norbert Treitz

Juegos con resistencias

¿Qué tienen que ver las resistencias eléctricas con descomponer un rectángulo en cuadrados de diferente tamaño?

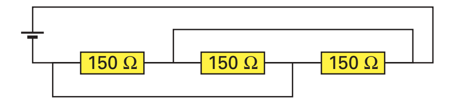
Para empezar, un problemilla. A este circuito eléctrico se le aplica tensión suficiente para que de las siete bombillas al menos algunas lleguen a encenderse. Pero, ¿brillarán A y B con la misma intensidad?



En corriente continua se denomina resistencia al cociente entre la tensión eléctrica en un segmento de conductor y la intensidad de la corriente que pasa por ahí. De acuerdo con la ley de Ohm, esa resistencia es independiente de la tensión, al menos para materiales de uso corriente y a temperatura constante. Por ello tiene sentido atribuirles a determinadas piezas de circuito eléctrico una resistencia eléctrica medida en ohm (Ω, volt entre ampère), imprimirla en la carcasa con números o colores codificados y denominar "resistencia" a la pieza. Si se conectan dos resistencias —las piezas— en serie, sus resistencias —los valores— se sumarán, pues se sumarán las correspondientes tensiones. Si se las conecta en paralelo, se sumarán las intensidades de corriente y, por tanto, los valores inversos de las resistencias (las "conductancias").

En el enredo de las siete bombillas el truco se esconde en su calentamiento desigual. En metales puros —vale también para los filamentos de las bombillas incandescentes—, la resistencia aumenta con la temperatura. A temperatura ambiente las cuatro bombillitas de abajo a la izquierda tienen en conjunto la misma resistencia que una sola; las condiciones en A y B serían, pues, las mismas. Pero una vez encendidas, cada una de esas cuatro no recibe ni la mitad de la tensión que cualquiera de las otras tres. Se calentarán menos. Y B recibirá una tensión mayor que A.

Otro acertijo: si se conectan en serie tres resistencias cada una de 150 Ω, se obtiene la suma de 450 Ω. Pero aquí se han soldado dos puentes más.



El problema se simplificará mucho si nos preguntamos qué nodos tienen el mismo potencial. (Despreciaremos la resistencia de los cables; supondremos que es cero.) Comprenderemos entonces que las resistencias están dispuestas en paralelo y, por ello, sólo suman 50 Ω.

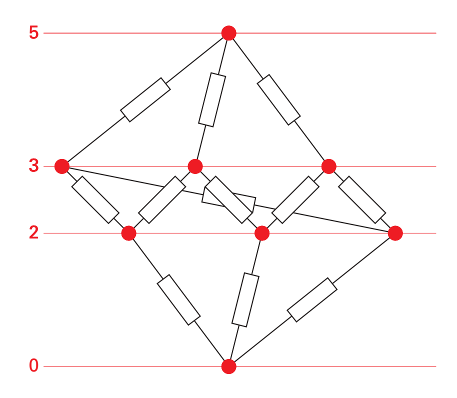
En sus "Juegos matemáticos" de *Scientific American* de diciembre de 1958, dedicados a los sólidos pitagóricos —o platónicos—, Martin Gardner planteó el siguiente problema: doce resistencias iguales, digamos que cada una de 1 Ω, están soldadas como las aristas de un cubo. ¿Cuál es la resistencia entre dos vértices opuestos? Muchos expertos efectuaron cálculos muy complicados.

¿Cómo se calcula la resistencia de una red poderosa de resistencias? Con frecuencia se pueden sustituir dos resistencias en serie o en paralelo por otra y seguir trabajando con la red así simplificada.

A veces se llega muy elegantemente a la solución mediante consideraciones de simetría. En el cubo de Gardner el quid —como él mismo explicó en enero de 1959— estribaba en reconocer que los tres vértices vecinos de uno de los dos vértices escogidos se hallaban conectados de la misma manera. Si se intercambian dos de esos vértices vecinos no cambia en absoluto la estructura de las conexiones, la "topología". Por ello, han de tener el mismo potencial. Obtendremos una conexión en serie de (1/3 + 1/6 + 1/3) Ω = 5/6 Ω.

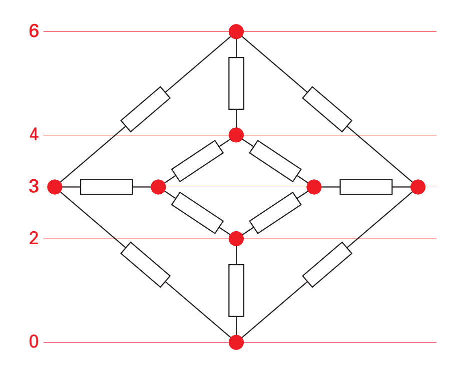
En lo que sigue tomaremos siempre 1 k Ω en vez de 1 Ω; tendremos menos problemas con el calentamiento y las resistencias de las conexiones.

En este gráfico el cubo se ha deformado de modo que los vértices tengan sus coordenadas *y* dispuestas a lo largo de una escala vertical de potenciales. Las coordenadas *x* parecen así repartidas más bien al azar.



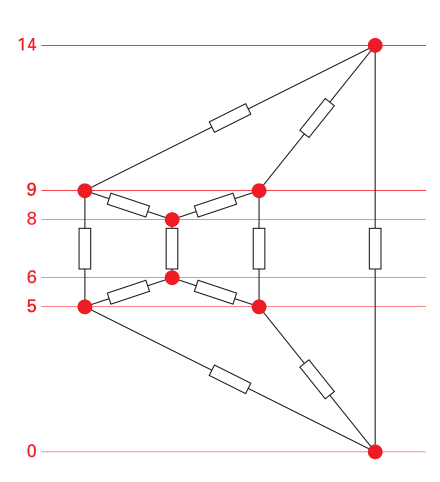
Spektrum der Wissenschaft, SEGUN NORBERT TREITZ

Cuando aplicamos la tensión entre dos vértices que sean segundos vecinos la gráfica es más clara:



Para 6 voltios fluyen 8 miliampère (mA).

Menos evidente es la resistencia entre dos vértices inmediatamente vecinos. Está claro que no llega a 1 k Ω porque las otras 11 aristas forman una conexión paralela de magnitud equiparable. Dibujando el cubo con una deformación adecuada, fijándose en vértices “con el mismo valor” y sustituyendo grupos de resistencias conectadas en paralelo y en serie paso a paso, se obtiene 7/12 k Ω. Para una tensión total de 14 volt, a cada resistencia le corresponderán valores enteros de la tensión en volt y de la intensidad de corriente en mA:



Para el cubo hay, pues, tres maneras de plantear el problema, para el octaedro (que también tiene 12 aristas) sólo dos y para el tetraedro sólo una, con 0,5 k Ω.

¿Qué haremos con redes más complejas de resistencias o que se compongan de resistencias diferentes? Plantearemos un sistema de ecuaciones lineales y lo resolveremos. Las incógnitas serán los potenciales en los nodos (de entrada, sólo se conocen los potenciales de los puntos en que se aplica la tensión). Las ecuaciones se obtienen a partir de las leyes de Kirchhoff: la suma de las intensidades de corriente —precedidas del signo adecuado— que fluyen hacia un nodo es igual a cero.

La resolución iterativa imita a la naturaleza. De entrada se le asigna a cada nodo, por valor inicial, un potencial arbitrario, con ayuda de un generador de números aleatorios. Empieza entonces un bucle de repeticiones en el que se calcula, para cada conductor, su diferencia de potencial y, a partir de ésta y la resistencia, la intensidad de corriente.

Cabe imaginar los nodos como minúsculas placas de condensador cuyo potencial es más elevado cuanta más carga fluya hacia ellos. Los potenciales de los nodos en los extremos de un conductor cambiarán ligeramente, de manera proporcional a la intensidad de la corriente. Esta variación no se tiene en cuenta de inmediato: se calculan las intensidades de corriente a través de todos los conductores a partir de los potenciales actuales, y sólo después se aplican a los potenciales las correcciones acumuladas. A los dos nodos conectados a los bornes se les reasignan sus valores originales y el programa vuelve a ejecutar el bucle.

Tras ejecutarlo unas decenas de miles de veces se obtienen en pocos segundos unos potenciales muy estables. En cada nodo —igual que en la naturaleza— se alcanza el equilibrio del flujo según Kirchhoff: se cumple la regla de los nodos. En cálculo numérico se llama a este método “procedimiento iterativo de Jacobi para sistemas de ecuaciones lineales”.

Cuando todas las resistencias valen 1 k Ω, los coeficientes del sistema de ecuaciones son números enteros bastante pequeños y sus soluciones, para una tensión aplicada de 1 volt, números racionales con un denominador no demasiado alto. Multiplicando la tensión con este denominador —al que se habrá llegado tanteando o calculando— se obtienen valores enteros para todos los potenciales de los nodos.

Rectángulos de hojalata

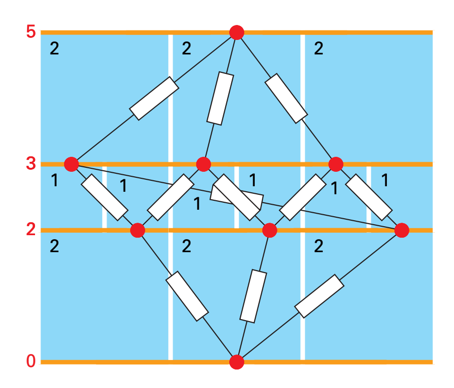
Imaginemos que disponemos de hojalata, de grosor pequeño y homogéneo y con resistencia específica no demasiado pequeña, y cable de cobre “ideal”, de resistencia pequeñísima. Soldemos el cable a dos aristas opuestas de un rectángulo recortado de la hojalata. La resistencia entre ellas será proporcional a la longitud del rectángulo y al inverso de su ancho, esto es, a la distancia entre los cables de cobre y al inverso de su longitud. Imaginaremos que el grosor y el material de la hojalata son tales que, para un trozo cuadrado —¡de cualquier tamaño!— tengamos justamente 1 k Ω.

Tomemos ahora un rectángulo grande de hojalata con los cables en los lados superior e inferior y apliquemos una tensión. La densidad de corriente en la hojalata será homogénea y el potencial, proporcional a la coordenada vertical. No cambiará nada si cortamos hendiduras verticales en la hojalata o si soldamos trozos horizontales de conductor ideal. También podríamos sustituir los vanos rectangulares, entre los conductores añadidos, por resistencias corrientes cuyos valores coincidan con la razón de los lados de esos rectángulos. Esta manera de descomponer un rectángulo equivaldrá a una estructura de resistencias sin cruces.

Hacia 1936, William T. Tutte y sus colegas C. A. B. Smith, H. Stone y R. L. Brooks, del Colegio Trinity de Cambridge, expusieron la relación entre la descomposición en cuadrados de una superficie y las técnicas de la corriente continua. Descubrieron así una conexión entre dos campos de las matemáticas recreativas completamente diferentes: por un lado, la descomposición de rectángulos, en especial de cuadrados, en cuadrados, y por otro el cálculo de redes compuestas de muchas resistencias iguales, entre ellas las que conforman esqueletos poliédricos. Martin Gardner trató este asunto en noviembre de 1958, un mes antes de plantear el problema del cubo [*véase también* “La cuadratura del cuadrado”, por Ian Stewart; INVESTIGACIÓN Y CIENCIA, septiembre de 1997].

Para el cubo, la descomposición de los rectángulos en cuadrados (correspondiente a la primera manera de plantear el problema) y la estructura de resistencias intercalada vienen dadas por la siguiente figura:

Spektrum der Wissenschaft, SEGÚN NORBERT TREITZ



Las barras naranja en los bordes de los cuadrados representan los hilos de contacto ideales; las blancas representan las rendijas aislantes, exagerando su grosor. Los números de las esquinas son los lados de los cuadrados y por tanto, al mismo tiempo, tensiones en volt e intensidades de corriente en mA (si nos quedamos con 1 k Ω para cada cuadrado individual).

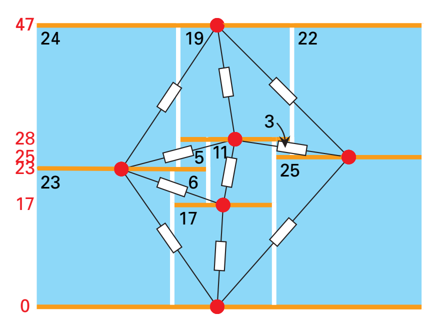
El gráfico del cubo se puede aplanar (representar en un plano sin cruces) dibujando una malla tan grande, que las demás se vean a través de ella y quepan en la misma. Pero si fijamos ciertos vértices, como cuando se conectan vértices opuestos, la superposición de mallas no se podrá evitar. En cualquier caso, cabrá imaginar que la red se dispone sobre la cubierta de un cilindro cuyo eje represente la escala de potenciales. Entonces, el rompecabezas de los cuadrados se formulará también sobre la cubierta finita del cilindro.

La cuadratura del cuadrado

Descomponer un cuadrado en muchos cuadrados iguales, como en un tablero de ajedrez, lo puede hacer cualquiera. Tampoco es demasiado difícil cortar un cuadrado en bandas de diferentes anchuras y de cada banda hacer una tira de cuadrados idénticos. Más interesantes son las descomposiciones de rectángulos —aún más, de un cuadrado— en muchos cuadrados diferentes.

Por desgracia, la simetría de una estructura, que nos simplifica el análisis, nos conducirá con seguridad a muchos cuadrados iguales. Para obtener descomposiciones menos aburridas hay que romper la simetría, por ejemplo quitando esta o aquella arista del esqueleto de aristas —muy simétrico— de un sólido pitagórico.

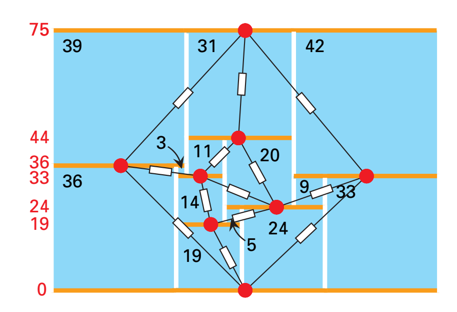
Al aplicar una tensión entre dos vértices vecinos de un octaedro sólo se encuentran cinco valores distintos de las tensiones e intensidades de corriente parciales en toda la estructura. Este número no cambia si se quita la conexión directa de los vértices vecinos. Pero si se retira otra arista, la figura se volverá mucho más irregular:



Tenemos una descomposición de un rectángulo de tamaño 47 por 65 en diez diferentes cuadrados. Recortándolo de un cartón ondulado, o mejor de una gomaespuma, conseguimos un rompecabezas (bastante fácil, por cierto).

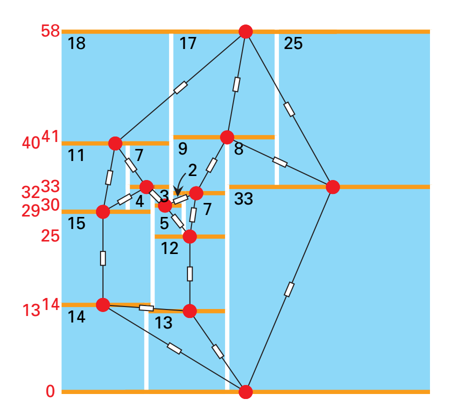
También se puede girar 90 grados. La correspondiente estructura de corrientes tiene siete peldaños de potencial e igualmente 10 resistencias. Si se le aplican 65 volt, fluirán entre los extremos 47 mA y por cada una de las resistencias (como antes, cada una de 1 k Ω), corrientes de número entero (en miliampère).

Sólo hay seis descomposiciones de un rectángulo en diez cuadrados distintos. Para menos cuadrados sólo hay —aparte del trivial rectángulo 1 × 1— dos soluciones: en 1925 Zbigniew Moron (1904-1971) encontró un rectángulo de 32 × 33 y posteriormente Tutte uno de 61 × 69, ambos con 9 cuadrados diferentes (*véase el artículo citado de Ian Stewart*). Para la descomposición encontrada por Brooks del rectángulo 112 × 75, hay dos posibilidades con los mismos 13 cuadrados distintos, a las que corresponden dos estructuras de resistencias diferentes. Una de ellas es el esqueleto de un cubo con una decimotercera resistencia adicional en una de las diagonales de una cara, que rompe la simetría:



El servicio postal alemán sacó un sello con la descomposición de un rectángulo de 176 × 177 en 11 cuadrados distintos, encontrada por Stone en 1936. Si se sabe que los más pequeños tienen una longitud de 9 y 16 unidades, los otros se infieren sin dificultad.

Como los sólidos de Arquímedes son menos simétricos, ofrecen más posibilidades. El tetraedro truncado, con una tensión aplicada entre vértices segundos vecinos, da lugar a la siguiente figura:



¿No es una pena que aparezca dos veces el siete?

Adaptando mi programa de ordenador, no sólo puedo generar estructuras de resistencias, sino también, a partir de ellas y de forma semiautomática, descomposiciones de rectángulos en no muchos cuadrados. Me parece tan interesante, al menos, como el sudoku.

La verdadera maña consiste en descomponer un cuadrado en muchos cuadrados distintos. A. J. W. Duijvestijn demostró en 1962 que, para descomposiciones perfectas, es decir, que no contengan ningún rectángulo compuesto de cuadrados, se necesitaban al menos 21 cuadrados. Pasarían 16 años antes de que hallase una solución de 21 cuadrados, y ni uno más, y demostrara además su univocidad.

Spektrum der Wissenschaft, SEGUN NORBERT TREITZ

JUEGOS MATEMÁTICOS



Juan M.R. Parrondo

Móviles y vectores

Termina el primer tiempo del partido en un gran estadio. En ese momento, miles de personas toman su teléfono móvil para llamar a casa o a algún amigo. Aunque en este tipo de situaciones las líneas suelen saturarse, son muchos los que consiguen llamar. ¿No se han preguntado nunca cómo es posible que la información proveniente de todos esos aparatitos pueda procesarse por separado? ¿Y cómo unas pocas antenas pueden distribuir a su vez información a todos ellos, de forma dirigida, específica y bastante rápida? Denle las gracias a las matemáticas. Más concretamente, a los vectores y matrices que tantos quebraderos de cabeza provocan en nuestros estudiantes de secundaria y de universidad.

En los comienzos de la telefonía móvil, la estrategia para que la estación base se pudiera comunicar con varios teléfonos al mismo tiempo fue asignar a cada uno de ellos una frecuencia o canal, como si se trataran de distintas cadenas radiofónicas. Este sistema de acceso múltiple por división de frecuencia (FDMA, acrónimo de Frequency Division Multiple Access) imperó en la telefonía móvil hasta los años noventa. Pero no utiliza el ancho de banda disponible de manera óptima. Entre otras razones porque, para evitar interferencias, cada antena o base no puede emplear todas las frecuencias disponibles. Las antenas de telefonía móvil se disponen formando una red aproximadamente hexagonal, como la que se muestra en la figura 1. Debido a posibles interferencias, una celda no puede utilizar las mismas frecuencias que sus vecinas, lo cual obliga a una cuidadosa planificación de las frecuencias asignadas a cada celda, como vemos en la figura 1 (*derecha*). Las celdas con distintas letras emplearán frecuencias diferentes; así, cada una de ellas hará uso sólo de un séptimo del ancho de banda disponible, con la pérdida consiguiente de capacidad para toda la red.

La segunda generación de telefonía móvil se basó en el sistema de acceso múltiple por división de tiempo, TDMA (de Time Division Multiple Access). En este sistema no sólo se divide el espectro de frecuencias, sino también el tiempo. En cada canal habla más de un usuario (3 en EE.UU. en canales de 30 kHz, y 8 en Europa y Asia en canales de 200 kHz) en intervalos de tiempo muy cortos, del orden de 30 o 40 milisegundos. Cada usuario utiliza este intervalo de tiempo de forma secuencial; el oído apenas nota las breves interrupciones, aunque hay una pérdida de calidad apreciable. Es un hecho reconocido que la segunda generación de telefonía móvil perseguía el aumento de capacidad, aunque fuera a costa de la calidad de la comunicación. Sin embargo, en este sistema, sigue siendo necesaria la planificación de frecuencias y continúa sin aprovecharse todo el ancho de banda disponible.

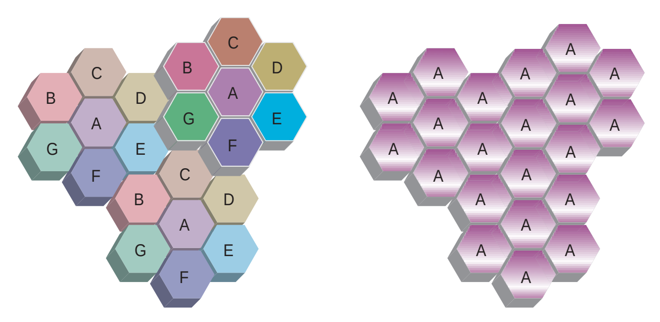
Por último, la 3G o tercera generación de telefonía utiliza el sistema de acceso múltiple por división de código, CDMA (de Code Division Multiple Access). Permite aprovechar, de manera muy ingeniosa, la misma frecuencia (con un ancho de banda mayor) para *todos* los usuarios que se encuentran en el área de alcance de una determinada antena o base. No se requiere ya ninguna planificación de frecuencias, como se muestra en la figura 1 (*izquierda*); cada celda de la red puede servirse, en principio, de todo el ancho de banda disponible.

El CDMA ha sido posible gracias a la digitalización de la señal que viaja de nuestros teléfonos a la base o antena más cercana (aunque esta digitalización ya se encontraba en el TDMA). El teléfono convierte nuestra voz en una señal digital compuesta por ceros y unos. La base también transmite al teléfono una señal digital y éste la convierte en el voltaje que finalmente reproduce el sonido en el altavoz. Sin embargo, para explicar el funcionamiento del CDMA es mejor recurrir a los valores +1 y –1, en vez de unos y ceros. En esta representación, que se denomina polar, +1 equivale a un uno y –1 equivale a un cero.

A cada teléfono que recibe o realiza una llamada en el área correspondiente a una base, ésta le asigna un número binario de varios dígitos, código que podemos representar como una serie de números, en este caso +1 o –1 (en matemáticas, un *vector*). Un código de cinco dígitos podría ser, por ejemplo:

$$\boldsymbol{b} = (+1, -1, -1, +1, -1)$$

El teléfono necesita enviar y recibir constantemente bits, es decir, ceros y unos. Sin embargo, en el sistema CDMA no los envía directamente, sino que transmite el código $\boldsymbol{b}$ si el bit que quiere mandar es uno y $-\boldsymbol{b}$ si el bit que quiere mandar es cero. Utilizando la representación polar con +1 y –1, podemos expresar tal prescripción de la siguiente forma matemática: si lo que el teléfono quiere enviar es s, deberá enviar $s\boldsymbol{b}$. Como el código $\boldsymbol{b}$ tiene determinada longitud, 5 dígitos en nuestro ejemplo, aunque 64 o más en las ejecuciones reales del CDMA,



1. Planificación de celdas y frecuencias en FDMA y TDMA (*izquierda*) y en CDMA (*derecha*).

necesitamos enviar todos estos bits para transmitir el bit s. Tamaño "derroche" nos permite identificar la fuente del bit entre todos los usuarios que se están comunicando con la base en un momento dado.

En la figura 2 podemos ver cómo funciona la codificación CDMA con el código $\boldsymbol{b}$ de cinco dígitos que hemos puesto de ejemplo. Según se aprecia en la figura, el número de bits que tiene que enviar el teléfono se multiplica por cinco. Normalmente un teléfono necesita enviar o recibir 9600 bits por segundo, para ello envía o recibe 1.228.800 bits por segundo utilizando distintos códigos, algunos propios del teléfono y otros propios de la base o asignados por ésta a una determinada conversación.

Cada teléfono o usuario tiene su propio código. Pese a ello, todos envían sus señales en la misma frecuencia. Por lo tanto, la base recibe la suma de todas ellas. Si están hablando simultáneamente 5 usuarios, por ejemplo, y si sus envíos están sincronizados, la señal que recibe la base en un intervalo de tiempo de longitud igual a la longitud de los códigos, es:

$$\boldsymbol{s} = s_1\boldsymbol{b}_1 + s_2\boldsymbol{b}_2 + s_3\boldsymbol{b}_3 + s_4\boldsymbol{b}_4 + s_5\boldsymbol{b}_5$$

en donde s_1, ..., s_5 son los bits que cada usuario quiere transmitir y $\boldsymbol{b}_1$, ..., $\boldsymbol{b}_5$ los correspondientes códigos. ¿Cómo puede la base inferir de esta suma la contribución de cada usuario? En este punto, el álgebra de vectores puede ayudarnos.

Una operación básica con vectores es el llamado *producto escalar*, que consiste en multiplicar las componentes de dos vectores una a una y sumar todos estos productos. Por ejemplo, el producto escalar de $\boldsymbol{a} = (1,5,4)$ y $\boldsymbol{b} = (3,0,2)$ es $\boldsymbol{a}\cdot\boldsymbol{b} = 1 \times 3 + 5 \times 0 + 4 \times \times 2 = 11$. Cuando el producto de dos vectores es nulo, decimos que son *ortogonales*. La denominación proviene de que, en espacios de dos y tres dimensiones, el producto escalar de dos vectores que forman un ángulo de 90 grados es nulo.

Sin embargo, la ortogonalidad es una propiedad más general. Así, en el caso de vectores cuyas componentes son sólo +1 y –1, que el producto escalar sea nulo es equivalente a que difieran en la mitad de sus componentes. Los vectores ortogonales nos van a permitir resolver el problema del CDMA. Supongamos que los códigos asignados a cada usuario son ortogonales entre sí; en tal caso, si multiplicamos escalarmente la señal recibida $\boldsymbol{s}$ por uno de los códigos, por ejemplo $\boldsymbol{b}_1$, el producto de dicho código por todos los demás será nulo. Tenemos entonces:

$$\boldsymbol{b}_1\cdot\boldsymbol{s} = s_1\boldsymbol{b}_1\cdot\boldsymbol{b}_1 = Ns_1$$

siendo N la longitud del código, ya que al multiplicar cada componente de $\boldsymbol{b}_1$ por sí misma el resultado es siempre 1 y la suma de todos esos productos conduce a $\boldsymbol{b}_1\cdot\boldsymbol{b}_1 = N$. Por lo tanto, si los códigos son ortogonales, la base puede recuperar la información enviada por cada usuario simplemente multiplicando escalarmente la señal recibida por el código correspondiente. El sistema de envío de información de la base a los teléfonos es similar. En este caso la base envía la suma $\boldsymbol{s}$ y son los teléfonos los que extraen de ella la información dirigida a cada uno.

La pregunta crucial para la capacidad de la red es entonces ésta: ¿cuántos códigos ortogonales existen y cómo podemos obtenerlos? El CDMA es ingenioso, pero no mágico: el número de códigos ortogonales es igual a la longitud de los mismos, o, formulado de modo más intuitivo y geométrico, el número máximo de vectores ortogonales en N dimensiones es precisamente N. Una forma particular de obtenerlos es a partir de las llamadas matrices de Hadamard-Walsh.

Se comienza con la matriz:

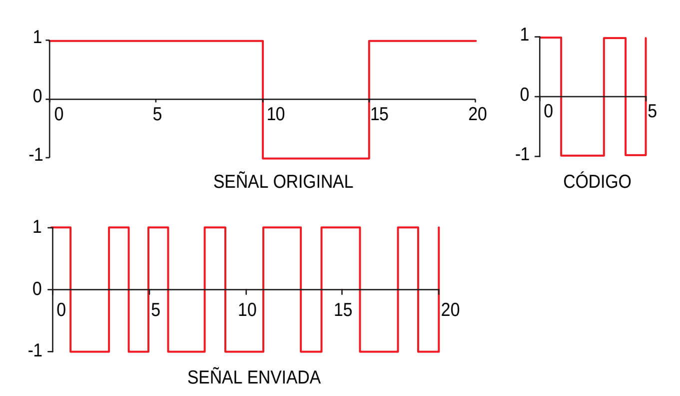
$$H_1 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

y se genera una nueva matriz mediante el siguiente algoritmo:

$$H_2 = \begin{pmatrix} H_1 & H_1 \\ H_1 & -H_1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}$$

Continuando de forma similar, se obtienen matrices en donde todas las columnas son vectores ortogonales entre sí. En la telefonía actual se utilizan estos códigos de Walsh con una longitud de 64 bits (aunque se extenderán a 256 en el próximo estándar de CDMA).

El sistema CDMA completo es bastante más complicado de lo que hemos esbozado aquí. Los códigos de Walsh se aplican a la transmisión base-teléfonos; para el envío de información de los teléfonos a la base se recurre a otro tipo de códigos que son sólo aproximadamente ortogonales. Si lanzamos dos monedas muchas veces, en más o menos la mitad de las tiradas ambas presentarán el mismo resultado. Por ello, dos vectores formados por +1 y –1 aleatorios tienden a ser ortogonales, lo cual permite construir códigos aproximadamente ortogonales mediante algoritmos muy simples que generan bits pseudoaleatorios. El mes próximo hablaremos de ellos.



2. La codificación de la señal en el sistema CDMA. La señal original se multiplica por el código del usuario y se envía a la base el resultado.

Mark Fischetti

RECONOCIMIENTO DE CARACTERES

Escritos

Escanea un libro para importar su contenido en un procesador de textos. Graba en una hoja de cálculo lo que se escribe a mano en la pantalla de un asistente personal digital (PDA). Descifra los garabatos de un formulario o el código postal de un sobre. En todos esos casos, un programa traduce los caracteres mecanografiados o manuscritos a un texto digital que puede editarse, enviarse por correo electrónico, almacenarse o usarse para decirle a una máquina de alta velocidad hacia dónde dirigir una carta.

Se denominaron al principio programas de reconocimiento óptico de caracteres; hoy esa expresión se refiere sólo al reconocimiento de textos tipografiados. El análisis de la escritura a mano en letra de imprenta o en cursiva se llama reconocimiento inteligente de caracteres. Independientemente de la etiqueta, ambos programas se basan en el mismo tipo de algoritmo para estimar los rasgos de una mancha de tinta; luego, comparan esos rasgos con modelos matemáticos para determinar con qué letra o número guarda mayor semejanza.

La determinación de caracteres escritos sobre un PDA constituye quizá la tarea más fácil, ya que el movimiento del estilo lo siguen sin dificultad la pluma o la pantalla. El análisis de un texto tipografiado o escrito a mano en una hoja impresa resulta más arduo, puesto que debe extraerse una señal de una imagen estática, limpiarla y luego distinguir las letras. Según David Doermann, del Laboratorio de Procesado del Lenguaje y de Medios de la Universidad de Maryland, la traducción de la cursiva libre es la que entraña mayor complejidad; los programas comerciales no ofrecen esa opción.

Años atrás, las técnicas de reconocimiento de caracteres se basaban en uno o dos algoritmos que comparaban patrones sencillos. Los programas actuales hacen uso de múltiples algoritmos; luego ponderan los resultados, o los someten a escrutinio, para tomar una decisión. Además, los programas antiguos tardaban varios minutos en convertir una página; los nuevos, en cambio, lo hacen en segundos, afirma Allan Stratton, directivo de Nunace Communications, fabricante de OmniPage, un programa de conversión de documentos.

Se están llevando a cabo investigaciones para la aplicación de esos programas a caracteres árabes, japoneses y chinos, cuyos trazos pueden representar palabras enteras, no sólo letras. También se están empezando a incluir programas de reconocimiento en las cámaras digitales y en los teléfonos celulares.

1. UN PROGRAMA QUE OCULTA LOS TRAZOS fuera de lugar reconoce el tipo de letra, enmarca cada palabra en un retículo y la segmenta en letras. Un algoritmo analiza luego los rasgos de cada letra (a la derecha se muestran los más comunes). En cada caso, el programa contrasta el vector de rasgos (un conjunto de números) con una tabla de valores que corresponden a posibles caracteres, signos ortográficos o trazos confusos (borrones). Para determinar la letra definitiva, ese programa "fuera de línea" pondera los resultados de distintos vectores de rasgos. Un posprocesado reúne los caracteres en palabras en un procesador de textos.



LAS PALABRAS se segmentan en letras.

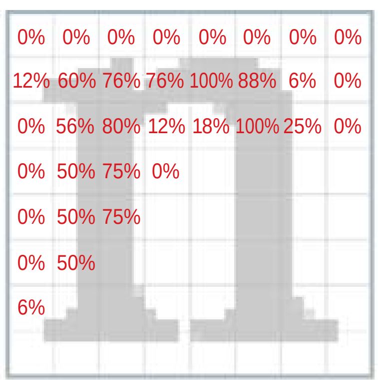
UN ESCANER capta una página de texto mecanografiado.

¿SABIA USTED QUE...?

➤ UNA NOVEDAD RELATIVA: Ya en 1929 se registraron las primeras patentes de sistemas de reconocimiento óptico de caracteres en Europa y América. La Agencia de Seguridad de las Fuerzas Armadas de los EE.UU. ensayó la técnica a principios del decenio de los cincuenta para automatizar la resolución de criptogramas. En 1965, el Servicio Postal estadounidense empezó con el escaneo electrónico de códigos postales; hoy la Oficina del Censo procesa por ese método millones de formularios. A principios del decenio de los noventa, el PDA de Apple Newton y el portátil ThinkPad de IBM acercaron al público el reconocimiento de la escritura a mano, aunque con escaso éxito comercial.

➤ BASES DE DATOS PARTICULARES, NO: Quienes, al comprar con tarjeta de crédito, firman sobre una pantalla electrónica junto a una caja registradora quizá crean que se analiza su firma. No es lo habitual. Según David Doerman, de la Universidad de Maryland, la mayoría de los comerciantes se limitan a guardar la imagen entera para ahorrar papel. La próxima vez, sugiere, trace algo completamente distinto y compruebe la reacción del empleado o la máquina.

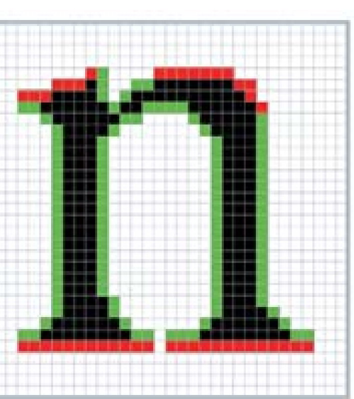
➤ CLASIFICACION: Los algoritmos que analizan los rasgos de los caracteres introducen sus conclusiones en programas clasificadores que las contrastan con tablas de referencia. Uno de los clasificadores comercializados es NN ("Nearest Neighbour", el semejante más parecido); compara los valores con todas las entradas de la tabla para determinar el parecido más estrecho. Otro clasificador es el HMM ("Hidden Markov Model", modelo oculto de Markov), que se aplicó primero al reconocimiento del habla; estima la distribución de probabilidad de las variables que conforman cada rasgo.



RASGO:
Color

ALGORITMO:
Encaje de matrices

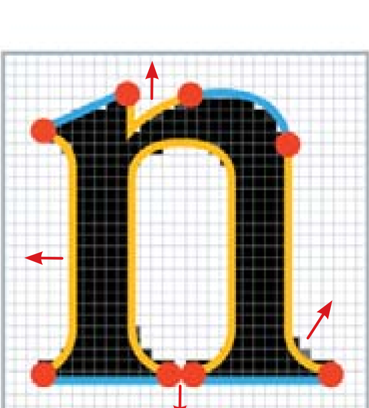
METODO:
Cálculo de la proporción de píxeles negros y blancos en cada celda



RASGO:
Número y longitud de las curvas

ALGORITMO:
Delineación de contornos

METODO:
Medición de curvaturas



RASGO:
Sentido de las curvas

ALGORITMO:
Delineación de contornos

METODO:
Ubicación de puntos extremos y de inflexión

DETERMINACION de los rasgos de una letra

2. DE RECONOCER LA ESCRITURA A MANO se encarga, casi en tiempo real, un programa "en línea". Cada trazo se define como el conjunto de puntos que hay entre un contacto de la punta de la pluma (pluma-abajo) y la siguiente elevación de la pluma (pluma-arriba). Dado que los puntos y los trazos se registran en función del tiempo, los algoritmos traducen esos datos en una cadena de números; contrastan luego esa cadena con unas tablas de referencia que indican el carácter más probable.

Pluma abajo
Pluma arriba
● Punta
■ Trazo
■ Tiempo

EL MOVIMIENTO vertical y horizontal del estilo es registrado por la pluma o la pantalla.

SW INFOGRAPHICS. FUENTES: NUANCE IMAGING *(letra n)*; VISION OBJECTS *("ink" manuscrito)*

Renacimiento

Realidades y símbolos

LE CONCEPT DE SEMENCE DANS LES THÉORIES DE LA MATIÈRE À LA RENAISSANCE. DE MARSILE FICIN À PIERRE GASSENDI, por Hiro Hirai. Brepols; Turnhout, 2005.

GIROLAMO FRACASTORO. FRA MEDICINA, FILOSOFIA E SCIENZE DELLA NATURA. Dirigido por Alessandro Pastore y Enrico Peruzzi. Leo S. Olschki; Florencia, 2006.

HERMAPHRODITES IN RENAISSANCE EUROPE, por Kathleen P. Long. Ashgate; Aldershot, 2006.

BEARING THE HEAVENS. TYCHO BRAHE AND THE ASTRONOMICAL COMMENTARY OF THE LATE SIXTEENTH CENTURY, por Adam Mosley. Cambridge University Press; Cambridge, 2007.

Distintiva del Renacimiento es la inyección de significados nuevos a términos vigentes desde la Antigüedad clásica. Así, el de semilla. El concepto renacentista de semilla, traído a primer plano por los neoplatónicos del Humanismo, constituye el nexo entre la doctrina aristotélico-escolástica de la forma sustancial y la teoría mecanicista de las “moléculas” de la modernidad (*Le concept de semence dans les théories de la matière à la Renaissance. De Marsile Ficin à Pierre Gassendi*). La noción trasciende el dominio de la embriología para abarcar los cimientos de la minería y la metalurgia. Hasta el punto de emparejar las “entrañas” de la Tierra con el seno materno. El iniciador, Marsilio Ficino (1433-1499), anudó una doble tradición, la doctrina estoica de los *logoi spermatikoi* —las “razones seminales” de san Agustín— y la idea lucreciana de las semillas de las cosas (*semina rerum*).

El recurso clásico estaba plenamente justificado. Anaxágoras denominó *sperma* al principio originante de las cosas naturales. De un punto seminal arrancaba el universo, según los pitagóricos. Y por principio seminal se entendió a lo largo de siglos la fuerza generatriz de la naturaleza. En la doctrina estoica de los *logoi spermatikoi* se unificaba la noción de fuerza creadora de la semilla de los seres vivos con la idea de semilla como origen de las cosas naturales y del universo. Cuando el neoplatónico Plotino (205-270) entra en la escena filosófica, sitúa los *logoi spermatikoi* en el alma, se trate del alma del mundo o de los particulares. En cualquier caso son inmateriales.

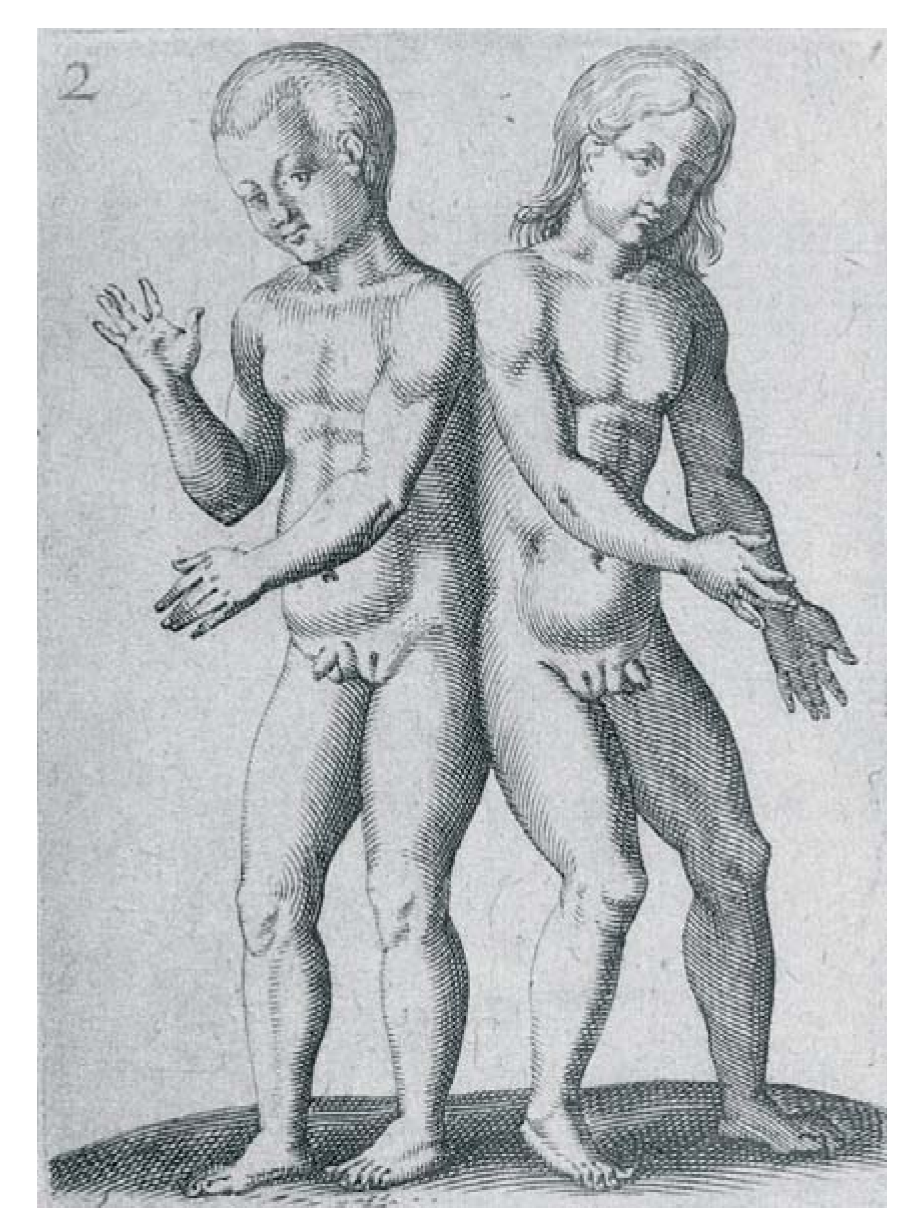
En la onda ficiniana, Girolamo Fracastoro (1478-1553), espejo de renacentistas en su amplitud de intereses (medicina, astronomía, cosmología, lógica, psicología y gnoseología). Nacido en Verona, hijo del conde Paolo Filippo Fracastoro y de Camilla Mascarelli, se formó en Padua, de cuya academia fue profesor. Publicó en 1530 su célebre poema *Syphilis sive morbus Gallicus*. Ocho años después apareció su tratado cosmológico *Homocentrica sive de stellis*. Luego, puso al día el *De contagione* en tres libros (1546), que venía precedido del *De sympathia et antipathia rerum*. La *editio princeps* de las *Opera omnia*, preparada por Paolo Ramusio, se imprimió en 1555 (*Girolamo Fracastoro. Fra Medicina, Filosofia e Scienze della Natura*).

Lo que en *Syphilis* es insinuación de una teoría de la simiente infecciosa, adquiere pleno desarrollo en el *De contagione*, donde aborda las semillas del contagio. *De contagione* expone la teoría, la tipología de las enfermedades contagiosas y sus tratamientos. Define el contagio como una infección que se experimenta en las “partículas mínimas e insensibles” que conforman el cuerpo humano. Y divide las enfermedades contagiosas en tres categorías: infecciones por contacto directo, contagiosas a través de los *fomites* y las transmitidas a distancia. Para explicar la primera categoría, infección por contacto, recurre al ejemplo de la putrefacción de los frutos puestos juntos. En su opinión, el principio y simiente de la putrefacción son las partículas mínimas, calientes y húmedas de los frutos. Las simientes provocan en el segundo cuerpo la putrefacción de las partículas que se había producido en el primero. Tales partículas formarían a su vez otras simientes. Mas para que la simiente pueda actuar, se requiere que encuentre una disposición en el cuerpo receptor afín a la del cuerpo del cual procede, de suerte que pueda adherirse y promover la putrefacción.

A los contagios transmitidos por los *fomites*, les aplica también la noción de simientes. Con el término “fomites” designa las cosas que no se corrompen por sí mismas, aunque sí conservan las simientes y afectan a otras. Se suponen receptores; las simientes se conservan largo tiempo en sus poros. Sobre los contagios transmitidos a distancia, declara que parecen un tipo especial de principio, de fuerza y tensión: están dotados de vigor para difundirse en un círculo, llegar muy lejos y penetrar muy deprisa. Imitan el movimiento circular de las “cualidades espirituales”. En breve, el contagio es el tránsito de una misma infección de un cuerpo a otro.

La noción de simpatía (y antipatía) constituye el fulcro de una teoría del contagio conciliable con la teoría humoral hipocrático-galénica. La doctrina de Fracastoro sobre la simpatía se inscribe en la reflexión hermético-neoplatónica sobre la concepción animada y armónica del cosmos. Desde los primeros capítulos del *De sympathia*, el *consensus universi* se identifica con el *mutuus*

Hermafroditas gemelos, macho y hembra. Tomado del libro de Bauhin, *De hermaphroditorum*... Cortesía de Division of Rare Manuscript Collections, Biblioteca Universitaria de Cornell.



nexus, que mantiene unidas las partes del cosmos según un orden armónico de perfección con la exclusión del vacío, factor de imperfección y de falta de conexión y comunicación entre los cuerpos.

La primera obra en prosa sobre el mal francés de Fracastoro, escrita en 1533, fue publicada en 1939; se trata de un manuscrito conservado en la Biblioteca Capitular de Verona. Se equivocó el traductor al verter la palabra *semina* de Fracastoro por *germen*, atribuyéndole el descubrimiento de los gérmenes. Tras los estudios de Vivian Nutton, nadie asimila ya los *seminaria* de Fracastoro y la teoría de los gérmenes de Koch.

Las enfermedades de transmisión sexual, categorizadas en la sífilis (el término tardaría en acuñarse), despertaron el interés de numerosos estudiosos. En verdad, la propia sexualidad fue una cuestión ampliamente debatida en el Renacimiento, con sus formas normales y aberrantes (*Hermaphrodites in Renaissance Europe*). En Francia, el hermafrodita se convirtió en imagen del reinado de Enrique III (rey de Francia de 1574 a 1589), atacado en panfletos políticos y en la novela de Thomas Artus *L'isle des hermaphrodites, nouvellement descouverte*, publicada en torno a 1605. En esta obra, el hermafrodita simboliza un país dividido en guerras civiles y su unificación potencial a través de la tolerancia pese a las diferencias.

Los panfletos políticos recurrían a la imagen del hermafrodita para representar el efecto subversivo que la Reforma ejercía sobre la unidad religiosa y política de Francia. Pero no sólo. El hermafrodita aparece también en los tratados alquimistas y médicos de Alemania y Suiza. En las obras de Ambroise Paré (1510-1615), cirujano de la familia real, y Jacques Duval (1555-1615), cirujano de Rouen, el hermafrodita representa la controversia mantenida entre cirujanos, que aprenden el arte mediante la observación del cuerpo vivo, y los médicos académicos de formación libresca. El hermafrodita nos habla entonces de la emergencia de la ciencia empírica y de la valoración de la práctica clínica por encima de la erudición.

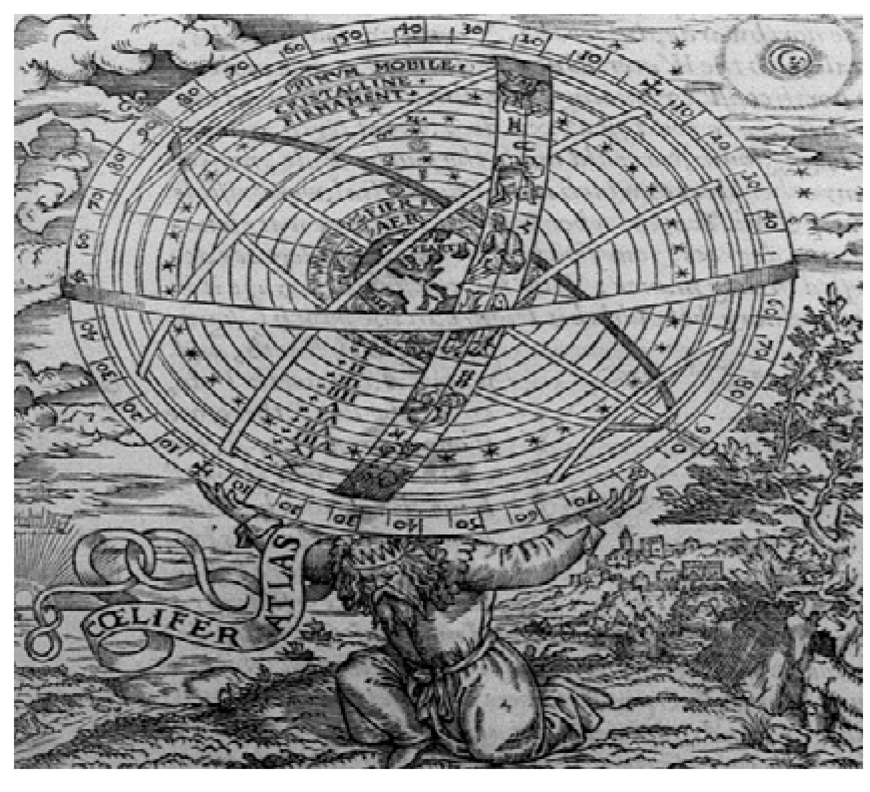
Sexo y género existen en una relación simbiótica fluida. Jacques Duval retrata a un Marin le Marcis feminizado tras un largo cautiverio y falta de actividad. Tesis que pone en cuestión incluso un sexo biológico fijo; el propio cuerpo parece condicionado por el ambiente. Se reconoce incluso una gradación, hasta el punto de que la intersexualidad se convierte en tema recurrente, casi obsesivo, sin excluir las obras alquimistas. Lo intersexual se celebra o se vitupera, se considera algo divino o algo diabólico. La naturaleza dual del hermafrodita, divino y monstruoso a un tiempo, arrancaba de tradiciones filosóficas y religiosas clásicas. El papel central del hermafrodita en el primitivo pensamiento gnóstico, como creador del universo, deviene referencia del hermafroditismo divino.

Al aproximarse el siglo XVI a su fin se incrementaron los tratados que giraban en torno a la anatomía masculina y femenina. Merecen destacarse el *Des monstres et prodiges*, de Paré; el *De hermaphroditorum monstrosorumque partuum natura*, de Bauhin, y *Des hermaphrodites et accouchements des femmes*, de Duval. Tales volúmenes reflejan una evolución gradual del pensamiento en lo concerniente al cuerpo, sexo, sexualidad y actitudes culturales en torno al género. Admitían por lo común cuatro categorías: varón con características

de mujer; mujer con características de varón; ambos; ninguno.

Tomemos el de Bauhin por muestra, *De hermaphroditarum* (1600). Ofrece la prueba más completa del estatuto cultural del hermafrodita en el Renacimiento tardío. En sus 25 páginas de obras citadas en su tratado, Bauhin aporta la fuente bibliográfica más completa. Docente de anatomía y botánica en Basilea y fascinado por la cuestión de la diferencia sexual, escribió otro tratado sobre los rasgos anatómicos diferenciadores entre varón y mujer: *Institutiones anatomicae corporis virilis et muliebris historiam exhibentes*. Concede al padre y a la madre un papel activo en la procreación. Se aleja así de la postura clásica que otorgaba a la madre una función pasiva. Para Bauhin el hermafrodita es *semivir*, es decir, su feminidad se expresa sólo como una carencia. En su opinión, el hermafrodita varón es la persona con apariencia de mujer (vulva), pero que sólo puede procrear como varón y carece de los órganos femeninos internos. De manera similar la hermafrodita hembra puede presentar la apariencia de un pene o testículos, pero no puede funcionar como varón. El hermafrodita neutro presenta signos de ambos sexos, pero no puede funcionar como ninguno. El cuarto tipo de hermafrodita tiene los órganos sexuales completos del varón y de la mujer.

Otro símbolo de circulación común en el Renacimiento fue el del atlante (*Bearing the Heavens. Tycho Brahe and the Astronomical Commentary of the Late Sixteenth Century*), del que fue notable impulsor Tycho Brahe (1546-1601). Llegó a la astronomía fascinado por la precisión matemática de las predicciones de los eclipses. Su éxito primero coincidió con la aparición de una nueva estrella brillante en 1572: demostró en *De nova stella* que esa "nova" se hallaba allende la Luna, es decir, en el dominio de las estrellas fijas e incorruptibles y los planetas. En Hven, donación de su mecenas Federico II de Dinamarca, construyó Uranibog, dotado de gabinete de instrumentos ópticos, observatorio y laboratorio alquimista. E imprenta. Sus observaciones tenaces del Sol, la Luna, los planetas y las estrellas le llevaron a revisar y mejorar hipótesis astronómicas y confeccionar catálogos que superarían los conocidos hasta entonces. Aunque sus resultados relativos a la medición de distancias arruinaban la disposición geocéntrica de los planetas, no adoptó el sistema copernicano, sino que propuso un sistema geo-heliocéntrico alternativo; el Sol y la Luna giraban en torno a una Tierra fija central, mientras que los planetas dibujaban órbitas alrededor del Sol.



El titán Atlas portador del firmamento, representado por la esfera armilar. El grabado es de la obra de W. Cuningham, *The Cosmographicall Glasse* (Londres, 1559). Cortesía de la Biblioteca Whipple, Universidad de Cambridge.

En 1596 publicó una selección de su propia correspondencia, *Epistolarum astronomicarum liber primus*. (Primero y único en vida del autor.) Predominan las cartas intercambiadas con el landgrave Wilhelm IV de Hesse-Kassel y con el *mathematicus* de su corte Christoph Rothmann. Movido hacia el estudio de los cielos tras la lectura de la obra de Pedro Apiano *Astronomicum Caesareum*, Wilhelm IV había sido educado en matemática por un hijo de Gerard Mercator. Estableció lo que puede ser considerado el primer observatorio de Europa occidental, en Kassel, y él mismo se entregó a la observación. En la epístola dedicatoria y en el poema que cerraba el libro, Tycho comparaba al landgrave con Atlas, el titán que portaba a sus espaldas el cosmos entero.

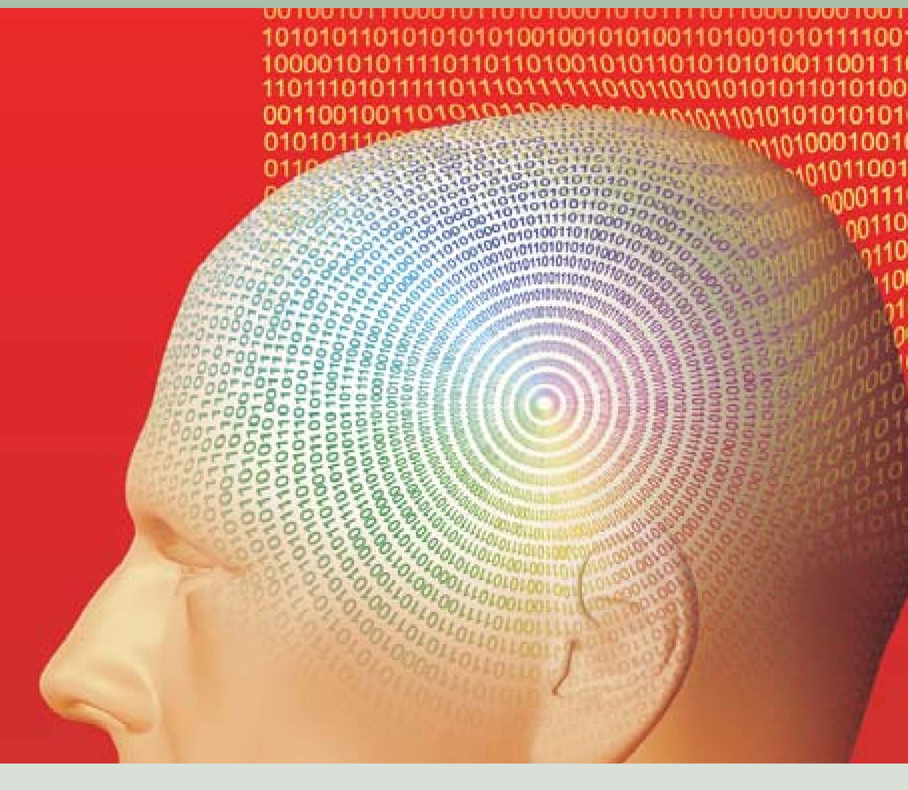
Las representaciones iconográficas del titán solían combinar dos formas de esquemas celestes. Una, la mostrada por los círculos concéntricos que indicaban las diez esferas del firmamento y las divisiones elementales del mundo sublunar: el *primum mobile*, los cielos cristalinos, el firmamento de las estrellas fijas, Saturno, Júpiter, Marte, el Sol, Venus, Mercurio, la Luna, Fuego, Aire, Agua, Tierra. La segunda forma de esquema del firmamento busca la idea de tridimensionalidad en la partición de la esfera celeste: los trópicos de Capricornio y Cáncer, el ecuador, los círculos polares, la línea de la eclíptica (la trayectoria aparente del Sol a lo largo del año en torno a la Tierra) y el Zodiaco que lo ciñe. El mito sirve de metáfora para la transmisión de teorías, datos y técnicas astronómicas. Adam Mosley considera cuatro modos principales de comunicación astronómica a comienzos de la era moderna: intercambio epistolar, producción editorial, manufactura de instrumentos y movilidad de los astrónomos.

—LUIS ALONSO



EL CODIGO DE LA MEMORIA, por Joe Z. Tsien

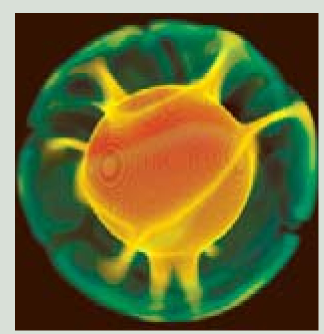
Cada vez nos hallamos más cerca de desentrañar el código de la memoria. Este se aplicaría al diseño de computadoras y robots más inteligentes.





EVOLUCION DE LOS FELINOS, por Stephen J. O'Brien y Warren E. Johnson

Las huellas genómicas en el ADN de los felinos salvajes de todo el mundo esclarecen el árbol genealógico de esta familia y revelan sus migraciones en el pasado.



KELVIN, PERRY Y LA EDAD DE LA TIERRA, por Philip C. England, Peter Molnar y Frank M. Richter

Si los científicos hubieran tenido en cuenta a uno de los críticos coetáneos de Kelvin, la teoría de la deriva continental podría haberse aceptado varios decenios antes.

CAMARAS OCULTAS EN LA PIRAMIDE DEL SOL, por R. Alfaro, E. Belmont Moreno, A. Cervantes, V. Grabski, J. M. López Robles, L. Manzanilla, A. Martínez Dávalos, M. Moreno y A. Menchaca Rocha

El descubrimiento de un túnel bajo la pirámide del Sol de Teotihuacan facilita la búsqueda de cavidades ocultas en este monumento arqueológico mediante la detección de rayos cósmicos.

UNA TIERRA SIN HUMANOS, entrevista con Alan Weisman

Un nuevo enfoque del impacto ambiental antropogénico se basa en el planteamiento de un mundo en el que la humanidad hubiese desaparecido.

